

## SEQUENCE LISTING

<110> Schupp, Thomas  
Ligon, James  
Molnar, Istvan  
Zirkle, Ross  
Cyr, Devon  
Goerlach, Joern

<120> GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

<130> 4-30582A

<140>

<141>

<160> 30

<170> PatentIn Ver. 2.0

<210> 1

<211> 68750

<212> DNA

<213> Sorangium cellulosum

<400> 1

```

aagcttcgct cgacgccctc ttgcgccgcg ccacctctgc ccgtgtgctc gatgatggcc 60
acggccgggc caccgagcgg catgtgctcg ccgaggcgcg cgggatcgag gacctccgcg 120
ccctccgaga gcacctccgc atccaggaag gggggccgctc ctttcaactgc atgtgcctcg 180
gcgacctgac ggtggagctc ctgcgcgacg accagccccct cgcgtccatc agcttccacc 240
atgcccgcag cctgaggcac cccgactgga cctcggaagc gatgctcgtc gacggccccg 300
cgctcgtccg gtggctcgcc gcgcgcggcg cgccgggtcc cctccgcgag tacgaagagg 360
agcgcgagcg agcccgaacc gcgcaggagg cgaggcgctt gtggctcgcg gccgcgcccg 420
cctgcttcgc gcccgatctg ccccgcttcg aggacgacgc caacgggctg ccgctcggcc 480
cgatgtcgcc tgaagtccgc gaggccgagc ggcgcctccg cgctcgtac gcgactcctg 540
agctcgccctg tgccgcgctg ctgcctggc tcgggacggg cgccgggtccc tggtcgggat 600
atcccgcccta cgagatgctg ccagagaatc tgctcctcgg gtttggcctc ccgaccgca 660
tcgcgcggcg ctccgcgccc ggcacatcgg aggcgctct ccgcggcgca gcgcggctgt 720
tcgcctcctg ggaggtcgta tcgagcaaga agagccagct cggcaacatc cccgaagccc 780
tgtgggagcg gctccggacg atcgctccgc cgatgggcaa tgccgacaac ctctctcgct 840
tcgagcgcg cgaggcgatc gcggcgagg tgccgcgctt gcgcgcacag ccggcgccct 900
tcgcggcggg cgccggcctg gcggctcgct gggctctctc gagcggccgg ctctcggggc 960
tcgtgaccga cggagacgca ttgtactccg gcgacggcaa cgacatcgtc atgttccaac 1020
ccggccggat ctgcgcggtc gtgctgctcg ccggaaccga tcccttcttc gagctcgcac 1080
cgccctcag ccagatgctc ttgcctcgcg acgccaacgc gggcaccatc tccaaggtcc 1140
tgacggaagg cagccccctc atcgatgagg caagaaacca ggccgcgacc atgagcctcg 1200
tccacgctcg cgggttcacg gcgtgggtca accaggccat ggtgcccagc cccgagcggg 1260
gcgcgccctt cgtcgtccag cgctcgacca tcatggaatt cgagcaccac acgcctcgct 1320
gtctccacga gcccgcgggc agcgctttct ccctcgctg cgacgaggag cacctctact 1380
gggtgcgagc ttccggtggc cggctcgagc tatggcgcca cccgcaccac cgccccggcg 1440
ccccgagccg cttcgcgta ctcggcgagc accccattgc ggccgacctg taccctcgc 1500
tcacctcaa tgcgaccac gtgctgtggg ccgaccctga tcgcagggcc atctcgggg 1560
tcgacaagcg caccggcgta gagcccatcg tcctcgcgga gacgcgccat ccccgggcgc 1620
acgtcgtgtc cgaggaccgg gacatcttcg cgcttaccgg acagcccagc tcccgcgact 1680
ggcacgtcga gcacatccgc tccggcgctt ggccgactac cagcgccagc 1740
tatgggacgg cctgacatg gtgctcaatc ggccggccct cttcttcacg acgaacgacc 1800
gcacctgac gctcgcggc agctgacatc gctcgacgcc gggccgctca tcgagggcgc 1860
ccggaccgag ctggcgaccc gccgctggcg ggccgcagct catgccgatt cgggtggcgac 1920
gtagacgctg cgccagaaac gctcgagagc ccccagaaac aggaagccgg cggattgtgt 1980
catcacgac ccgatcagct cgcggcccg atcattgatc caggacgtcc cgaacccgcc 2040
gtcccacca tagcgccgg gcacctccga gaccgcgtcc ggccgctga ccacggccat 2100
cccataacc cagccgtgcg tctcgaagaa gcccgggaaa aacgaggacg ccgccttctg 2160
ggccggcggt aggtgatcg ccgtcatctc gcgcaccgag gcggcgctca agagccggcg 2220
gccctcgtgc acaccgccgt tcatgagcat gcgcgcgaac aggaggtagt cgtccaccgt 2280
cgacacgagc ccggcggcgc ccgaaggaa cgccggcggg ctggcatagg cgctctcggc 2340

```

cccgtcgcga	tccatgcgcg	tcttctcccc	cgtctgctcg	tcggtgaagt	aaccgcagcc	2400
cgcaaacga	gcgagcttgt	ccgcccggac	gtgaaagtgc	gtgtcccga	tcccagcgg	2460
cgcgaggatg	cgctcgcgca	cgaacgcata	gaagccctgg	tcggccgcgc	gccccacgag	2520
cacccctgc	accaggtctc	ccgtgttgta	catccactgc	gccccggct	gatgcatgag	2580
cggcagcgtc	ccgagccgcc	ggatccactc	gtctggcccg	tcggcgctca	tcggcaccgg	2640
ctgcgcgttg	acgagcccga	gctcgtcgat	ggcccgctgg	atcggcgacg	atgctcgaa	2700
cgagattccg	aagcccatcg	tgaacgtcat	caggtcgcgc	accgtgatcg	gccgctccgc	2760
gggcaccgtc	tcgtcgatcg	gaccatcgat	gcgcgccagc	accttccggt	tcgcgagctc	2820
cggaacccat	cggtcgacgg	gggagtcgag	gtcagacttg	ccttctctga	cgagcatcat	2880
caccgcgctc	gcggtgaccg	ccttcgtcat	cgaggcgatc	cggaagatcg	tgtcccgcgg	2940
catgggcgcg	ctgccgcgca	gctcgggtcac	gcccaccgcg	tccacgtgca	cgctcgtccc	3000
gcgcgcgacc	agccagaccg	ctcccggcat	ctgccccgcc	gccacctccg	ccgccatcac	3060
ctcgcgcgcg	ggcgccagcg	cgccggcccc	cgcgtcctgc	cctggctgcc	cctcctcctc	3120
ggccccaccc	aacgcgcacc	ccggcgccgc	cacgtgatc	aaagctccca	taaactcccg	3180
ccttctcatg	accgtcgatg	cctctccgag	cgggggcgcc	tgccctgcc	gagagcactg	3240
actgcccgcg	cccgaaaaaa	tcatcggtgc	cccgtcacga	tcgcgcgcgg	gcgtggctcc	3300
gcccggccgc	ccgctcgggc	gcccggccct	ggacgagcaa	agctcgcccc	cccgcgctca	3360
gcacgcgcgt	tgccatgtcc	ggcctgcacc	cacaccgagg	agccaccac	cctgatgcac	3420
ggcctcaccg	agcggcaggt	cctgctctcg	ctcgtcacc	tcgcgctcat	cctcgtgacc	3480
gcgcgcgcct	ccggcgagct	cgcgcgcggg	ctgcgccagc	ccgaggtgct	cggggagctc	3540
ttcggcgggc	tcgtgctggg	ccctcctcgc	gtcggcgcg	tcgcgcgcgg	gttccatcga	3600
gcccctcttc	aggagccggc	ggtcgggggtc	gtgctctcgg	gcattctcctg	gataggcgcg	3660
ctcctcctgc	tgctgatggc	gggcatcgag	gtcgacgtgg	gcattctcgcg	caaggaggcg	3720
cgccccgggg	cgctctcggc	gctcggcgcg	atcgcgcgcc	cgctcgcggc	gggcgcgcgc	3780
ttctcggcgc	tcgtgctcga	tcggcccttc	ccgagcggcc	tcttctcgcg	gatcgtgctc	3840
tcggtgacgg	cggtcagcgt	gatcgcaag	gtgctgatcg	agcgcgagtc	gatgcgcgcg	3900
agctatgcgc	aggtgacgct	cgcgccgggg	gtggtcagcg	aggtcgctgc	ctgggtgctc	3960
gtcgcgatga	cgctcgtcag	ctacggcgcg	tcgcccgcgc	tgggcggtcgc	ccggagcgcg	4020
ctcctggcga	cggtattctt	gctgttcattg	gtgctcgtcg	ggcgccggct	caccacctc	4080
gcgatgcgct	gggtggccga	cgcgacgcgc	gtctccaagg	gacaggtgtc	gctcgtcttc	4140
gtcctcacgt	tcctggccgc	ggcgctgacg	cagcggctcg	gcctgcaccc	gctgctcggc	4200
gcgttcgcgc	tcggcggtgt	gtcaacagc	gtcctcgcga	ccaaccgccc	tctcctcgac	4260
ggcgtgcaga	cgctcgtggc	gggcctcttc	gcgcctgtgt	tcttcgtcct	cgcgggcatg	4320
cgcgctcgacg	tgctcgagct	gcgcagcccg	ggcgctggg	ggacggctgc	gttgctgtg	4380
gcgaccgcga	cgggcgccga	ggctcgtccc	gcccgcctcg	gcgcgcggct	cgggcggtc	4440
aggggcagcg	aggcgccgct	cgtggcggtg	ggcctgaaca	tgaaggcgcg	cacggacctc	4500
atcgctcgca	tcgtcggcgt	cgagctcggg	ctcctctcca	acgaggctta	tacgatgtac	4560
gccgtcgtcg	cgctggtcac	ggtgaccgcc	tcacccgcgc	tcctcatctg	gctcgagaaa	4620
agggcgccctc	cgacgcagga	ggagtcgggt	cgctcgcagc	gcgaggaggc	cgcgaggcgc	4680
gcgtacatcc	ccggggtcga	gcggatcctc	gtcccgatcg	tgccgcacgc	cctgcccggg	4740
ttcgccacgg	acatcggtga	gagcatcgct	gcctccaagc	gaaagctcgg	cgagacgggt	4800
gacatcacgg	agctctccgt	ggagcagcag	gcgcccggcc	catcgcgcg	cgcgggggag	4860
gcgagccggg	ggctcgcgag	gtcggcgcg	cgcctccgcg	tcggcatctg	gcggcaagg	4920
cgcgagctgc	gcggctcgat	ccaggcgatc	ctgcgcgcct	cgcgggatca	cgatctgctc	4980
gtgatccggc	cgcgatcgcc	ggcgcgcgcg	cgcggaatgt	cgttcggtcg	cctgcaggac	5040
gcgatcgtcc	agcgggcccga	gtccaacgtg	ctcgtcgtgg	tgggcgaccc	tccggcgggc	5100
gagcgcgcct	ccgcgcggcg	gatcctcgtc	cgatcatcg	gcctcgagta	ctccttcgcc	5160
gccgccgatc	tcgcggccca	cgtggcgctg	gcgtgggacg	ccgagctcgt	gctgctcagc	5220
agcgcgcaga	ccgatccggg	cgcggtcgtc	tgccgcgatc	gcgagccatc	ccgggtgcgc	5280
gcggtggcgc	ggagcgtcgt	cgacgaggcg	gtcttccggg	ggcgccgggt	cggcgtgcgc	5340
gtctcgtcgc	gcgtgcacgt	gggcgcgcac	ccgagcgacg	agataacgcg	ggagctcgcg	5400
cgcgcgccctg	acgatctgct	cgtgctcgga	tgctacgacc	atgggcccgt	ggcccggtc	5460
tacctcgcca	gcacggtcga	gtcgggtggg	gtccggagcc	gggtgcgggt	cgcttgctc	5520
gtcgcgcatg	gagggactcg	agagcagggtg	aggtgaggct	tccaccgcgc	tcgcccgtga	5580
ggaagcgagc	gcccggctct	gccgacgatc	gtcactcccg	gtccgtgtag	gcgatcgtgc	5640
tgagcagcgc	gttctccgcc	tgacgcgagt	cgagccgggt	atgctgcacg	acgatggggg	5700
cgtccgattc	gatcacgctg	gcatagtccg	tatcgcgcg	gatcggtcgc	ggttcgggtca	5760
gatcgttgaa	ccggacgtgc	cgggtgcgcc	tcgctggaac	ggtcaccccg	taaggcccg	5820
cggggtcgcg	gtcgtgaaag	taaaccggtga	tgggcgacctg	cgctccccgg	tccgacgcat	5880
tcaacaggca	ggcgtctca	tggtcgtca	tctgcggctc	aggtccgttg	ctccccgctg	5940
ggatgtagcc	ctctgcgatt	gcacagcgcg	tccgcccgat	cggcttgctc	atgtgtctc	6000
cctcctggct	cctctttggc	agcctccctc	tgctgtccag	gagcgatggc	ctcttcgctc	6060
gacgcgctcg	gggatccatg	gctgaggatc	ctcgccgagc	gctccctgcc	gaccggcgcg	6120
ccgagcgccg	acgggctttg	aaagcgcgcg	accggccagc	ccggacgcgg	gcccagagag	6180
gacagtgggt	ccgccgtgaa	gcagagaggc	gatcgaggtg	gtgagatgaa	acacgtcgac	6240
acgggcccgc	gattcggccg	ccggataggg	cacacgctcg	gtcttctcgc	gagcatggcg	6300

ctcgccggct	gcgccgggtcc	gagcgagaaa	accgtgcagg	gcacgcggct	cgccccggc	6360
gccgatgcgc	gcgtcaccgc	cgacgtcgac	cccagcgccg	cgaccacgcg	gctggcggtg	6420
gacgtcgttc	acctctcgcc	gcccagcgcg	ctcgaggccg	gcagcgagcg	gttcgtcgtc	6480
tggcagcgtc	cgagccccga	gtccccgtgg	cgacgggtcg	gagtgctcga	ctacaatgct	6540
gacagccgaa	gaggcaagct	ggccgagacg	accgtgccgt	atgccaaactt	cgagctgctc	6600
atcacccgcg	agaagcagag	cagccctcag	tcgccatcgt	ctgccgccgt	catcgggccg	6660
acgtctgtcg	ggtgacatcg	cgctatcagc	agcgttgagc	ccgccagcag	gccccagggc	6720
cctgcctcga	tggccttccc	catcacccct	gcgcactcct	ccagcgacgg	ccgcgcagcg	6780
acggccgcgt	ccaagcaacc	gccgtgccgg	cgcgctcca	cgcgcgcgac	aggcgagcgt	6840
cctggcgcg	cctgcgcac	gctggaagga	tcggcgaggc	atggatagag	aatcgaggat	6900
cgcgatcttt	gttgccatcg	cagccaacgt	ggcgatcgcg	gcggccaagt	tcacgcgccg	6960
cgccgtgacc	ggcagctcgg	cgaggcggtt	gccgacttcg	gcggcgctcc	gcgcgtgctg	7020
ctctacgaca	acctcaagag	cgccgtcgtc	gagcgccacg	gcgacgcgat	ccgggtccac	7080
cccacgctgc	tggctctgtc	ggcgcattac	cgcttcgagc	cgcgccccgt	cgccgtcgcc	7140
cgcggcaacg	agaagggccg	cgctcagcgc	gccatcacgg	cgtggacgac	atggcgcgga	7200
aacgtcgctc	taaccgcccc	gcaatgtcat	gggaatggcc	ccttgaaatg	gcccccttag	7260
ggggctggcc	ggggctcgac	atatcgcgcg	atctccccgt	caattcccga	tggtaaaaga	7320
aaaatttgtc	atagatcgta	agctgtgata	gtggctctgtc	ttacgttgcg	tcttcgcac	7380
ctcgagcgag	ttctctcgga	taacttcaa	tttttcgag	gggggcttgg	tctctgggtc	7440
ctcaggaagc	ctgatcgggg	cgagctaatt	cccatccatt	tttttgaggc	tctgtctaaa	7500
gggattagat	cgagtgcgac	agttcttttg	cagtgcgcga	agaacctggg	cctcgaccgg	7560
aggacgatcg	acgtccgcga	gcgggtcagc	cgctgaggat	gtgcccgtcg	tggcggtatc	7620
tcccatcgag	cgcgcagccg	aagatccgat	tcgatcgtc	ggagcgagtt	gccgtctgcc	7680
cgggtggcg	atcgatctga	gcgggttctg	gacgctctc	gagggctcgc	gcgacaccgt	7740
cgggcgagtc	cccgcggaac	gctgggatgc	agcagcggtg	tttgatcccc	accccgatgc	7800
cccgggggaa	acgcccgtta	cgcgcgcatc	tttctgagc	gacgtagcct	gcttcgacgc	7860
ctccttcttc	ggcatctcgc	ctcgcgaaag	cgctcggtatg	gacctgcac	atcgactctt	7920
gctggagggtg	tgctgggagg	tcggtggaga	gcgcgcgac	gctccatcgg	cgctcgtcgg	7980
tacggaaacg	ggagtgttca	tcgggatcgg	cccgtccgaa	tatgaggccg	cgctgccgca	8040
agcgacggcg	tccgcagaga	tcgacgctca	tggcgggctg	gggacgatgc	ccagcgtcgg	8100
agcgggcccga	atctcgtatg	ccctcgggct	gcgagggccg	tgtgtcgcg	tgggacggc	8160
ctattcgtcc	tcgctggtgg	ccgttcatct	cgctggtcag	agcttgcgct	ccggggaatg	8220
ctccacggcc	ctggctgtgtg	gggtatcgct	gatgttgctg	ccgagcacc	tcgtgtggct	8280
ctcgaagacc	cgggcgctgg	ccagggacgg	tcgctgcaag	gcattttcgg	cggaggcccg	8340
tgggttcgga	cgaggcggaag	ggtgcgccgt	cggtgtcctc	aagcggctca	gtggagcccc	8400
cgcgagcggc	gatcggatat	tggcggtgat	tcgaggatcc	gcgatcaatc	acgacgggtg	8460
gagcagcgg	ctgaccgtgc	cgaacgggag	ctcccaagaa	atcgtgctga	aacgggccct	8520
ggcgagcgca	ggctgcgccg	cgtcttcggt	gggttatgtc	gaggcacacg	gcacggggc	8580
gacgcttggt	gaccccatcg	aaatccaagc	tctgaatcgg	gtatacggcc	tcgggcgaga	8640
tgtcgccacg	ccgctgctga	tcgggtcggt	gaagccaac	cttggccatc	ctgagtatgc	8700
gtcggggatc	actgggtctgc	tgaaggctct	cttgtccctt	cagcacgggc	agattcctgc	8760
gcacctccac	gcgcaggcgc	tgaacccccg	gatctcatgg	ggtgatcttc	ggctgaccgt	8820
cacgcgcgcc	cggacaccgt	ggccggactg	gaatacgccg	cgacgggcgg	gggtgagctc	8880
gttcggcatg	agcgggacca	acgcgcacgt	ggtgctggaa	gaggcgccgg	cggcgacgtg	8940
cacaccgcgg	gcgcgggagc	gaccggcaga	gctgctgggtg	ctgtcggcaa	ggaccgcgtc	9000
agccctggat	gcacaggcgg	cgcggtcgcg	cgaccatctg	gagacctacc	cttcgcagtg	9060
tctgggcatg	gtggcggttca	gtctggcgac	gacgcgcagc	gcgatggagc	accggctcgc	9120
ggtggcgggc	acgtcgaggg	aggggctcgg	ggcagccctg	gacgtgcgg	cgaggggaca	9180
gacgtcgccc	ggtgcgggtg	gcagtatcgc	cgattcctca	cgcggaagc	tcgctttct	9240
cttcaccgga	cagggggcgc	agacgctggg	catgggccgt	gggctgtacg	atgtatgggtc	9300
cgcgttccgc	gaggcggttcg	acctgtgcgt	gaggctgttc	aaccaggagc	tcgaccggcc	9360
gctccgcgag	gtgatgtggg	ccgaaccggc	cagcgtcgac	gccgcgctgc	tcgaccagac	9420
agccttcacc	cagccggcgc	tgttcacctt	cgaatatcgc	ctcgccgcgc	tgtggcggtc	9480
gtgggggtga	gagccggagt	tggtcgccgg	ccatagcatc	ggtgagctgg	tggctgcttg	9540
cgtggcgggc	gtgttctcgc	ttgaggacgc	ggtgttctctg	gtggctgcgc	gcggggcgct	9600
gatgcaggcg	ctgccggccg	gcggggcgat	ggtgtcgatc	gaggcgccgg	aggccgatgt	9660
ggctgctgcg	gtggcgccgc	acgcagcgtc	ggtgtcgatc	gcccgggcca	acgctccgga	9720
ccaggtggtc	atcgcgggcg	ccggggcaac	cgctgcagcg	atcgcgccgg	cgatggccgc	9780
gcgcggggcg	cgaaccaagg	cgctccacgt	ctcgcatcgc	ttccactcac	cgctcatggc	9840
cccgatgctg	gaggcggttcg	ggcggtgtgg	cgagtcgggtg	agctaccggc	ggccgctcat	9900
cgtcctggtc	agcaatctga	gcgggaaggc	ttgcacagac	gaggtgagct	cgccgggcta	9960
ttgggtgcgc	cacgcgcgag	aggtggtgcg	cttcgaggat	ggagtgaagg	cgctgcacgc	10020
ggccgggtgc	ggcaccttcg	tcgaggtcgg	tccgaaatcg	acgctgctcg	gcctgggtgc	10080
tgcctgcacg	ccggagccgc	ggccggcgctg	gctcgcacgc	tcgcgcgctg	ggcgtgacga	10140
gccggcgacc	gtgctcgagg	cgctcgccgg	gctctggggc	gtcggtggcc	tggctccttg	10200
ggccggcctc	ttccctcag	ggggggcgcg	ggtgccgctg	cccacgtacc	cttggcagcg	10260

cgagcgctac	tggatcgaca	cgaaagccga	cgacgcggcg	cgtggcgacc	gccgtgctcc	10320
gggagcgggt	cacgacgagg	tcgaggaggg	gggcgcggtg	cgcggcggcg	accggcgag	10380
cgctcggctc	gaccatccgc	cgcccagag	cggacgccgg	gagaaggctc	aggccgccgg	10440
cgaccgtccg	ttccggctcg	agatcgatga	gccaggcggtg	cttgatcacc	tcgtgcttcg	10500
ggtcacggag	cggcgcgcgc	ctggctcggg	cgaggtcgag	atcgccgtcg	acggcgccgg	10560
gctcagcttc	aatgatgtcc	agctcgcgct	gggcatgggtg	cccgcgcacc	tgccgggaaa	10620
gccaaccct	ccgctgctgc	tcggaggcga	gtgcgcgggg	cgcacgtcgc	ccgtgggcca	10680
gggcgtgaac	ggcctcgtgg	tgggccaacc	ggtcatcgcc	ctttcggcgg	gagcgtttgc	10740
taccacgtc	accacgtcgg	ctgcgctggt	gctgcctcgg	cctcaggcgc	tctcggcgat	10800
cgaggcggcc	gccatgcccc	tcgcgtacct	gacggcatgg	tacgcgctcg	acagaatagc	10860
ccgccttcag	ccgggggagc	gggtgctgat	ccatcgggcg	accggcgggg	tcggtctcgc	10920
cgcggtgcag	tgggcgcagc	acgtgggagc	cgaggtccat	gcgacggccg	gcacgcccga	10980
gaaacgcgcc	tacctggagt	cgctgggcgt	gcggtatgtg	agcgattccc	gctcggaccg	11040
gttcgtcgcc	gacgtgcgcg	cgtggacggg	cggcgaggga	gtagacgtcg	tgctcaactc	11100
gctctcgggc	gagctgatcg	acaagagttt	caatctcctg	cgatcgacac	gccggtttgt	11160
ggagctcggc	aagcgcgact	gttacgcgga	taaccagctc	gggctgcggc	cgttcctgcg	11220
caatctctcc	ttctcgtctg	tggatctccg	ggggatgatg	ctcgagcggc	cggcgcgggt	11280
ccgtgcgctc	ttggaggagc	tcctcggcct	gatcgcggca	ggcgtgttca	cccccccc	11340
catcgcgacg	ctcccagctg	cccgtgtcgc	cgatgcgttc	cggagcatgg	cgcaggcgca	11400
gcatcttggg	aagctcgtac	tcacgctggg	tgaccggag	gtccagatcc	gtattccaac	11460
ccacgcaggc	gccggccccg	ccaccgggga	tcgggacctg	ctcgacaggc	tcgcgtcagc	11520
tgcgcggccc	gcgcgcgcgg	cggcgttggg	ggcgttcctc	cgtacgcagg	tctcgcaggt	11580
gctgcgcacg	cccgaatatc	aggtcggcgc	ggaggcgctg	ttcaccgcgc	tcggcatgga	11640
ctcgtcatg	gccgtggagc	tgctgaaatc	tatcgaggcg	agcctcaagc	tgaagctgtc	11700
gacgacgttc	ctgtccacgt	cccccaatat	cgccttgttg	gccccaaaacc	tggtggatgc	11760
tctcgccaca	gctctctcct	tggagcgggt	ggcggcggag	aacctacggg	caggcggtgca	11820
aaacgacttc	gtctcatcgg	gcgcagatca	agactgggaa	atcattgccc	tatgacgac	11880
aatcagcttc	tgaacgagct	cgagcaccag	ggatatcaagc	tggcggccga	tggggagcgc	11940
tccagatac	aggcccccaa	gaacgcctcg	aaccgcaacc	tgctcgctcg	aatctccgag	12000
cacaaaagca	cgatcctgac	gatgctccgt	cagagactcc	ccgcagaatc	catcgtgccc	12060
gccccagccg	agcggcacgc	tcggtttcct	ctcacagaca	tccaagaatc	ctactggctg	12120
ggccggacag	gagcggttac	ggtcgccagc	gggattccag	cctatcgcca	atacgaactg	12180
acggatctcg	acgtgcccag	gctgagcgcc	gcttttcgga	aagtcgtcgc	gcggcacgac	12240
atgcttcggg	cccacacgct	gcccagacatg	atgcagggtga	tcgagcctaa	agtcgacgcc	12300
gacatcgaga	tcacgatctc	gcgcgggctc	gaccggagca	cacgggaagc	gaggctcgtg	12360
tcgttgccag	atgcgatgtc	gcaccgcac	tatgacaccg	agcgccctcc	gctctatcac	12420
gtcgtcgccg	ttcggctgga	cgagcggcaa	accgcgtcgc	tgctcagtat	cgatctcatt	12480
aacgttgacc	ttgctgcgct	gtccatcact	ttcaaggact	ggctcagctt	ctacgaagat	12540
cccagacact	ctctccctgt	cctggagctc	tcgtaccgcg	attatgtact	cgcgctggag	12600
tctcgcaaga	agtctgaggc	gcataacaga	tcgatggatt	actggaagcg	gcgcacgcc	12660
gagctccac	ctccgcgcac	gcttcgatg	aaggccgatc	catctaccct	gaaggagatc	12720
cgtttccggc	acacggagca	atggctgcgc	tcggactcct	ggggtcgatt	gaagcggtcg	12780
gtcggggagc	gcgggctgac	cccagcggcg	gtcatcctgg	ctgcattttc	cgagggtgatc	12840
gggcgctgga	gcgcgagccc	ccggtttacg	ctcaacataa	cgctcttcaa	ccggctcccc	12900
gtccatccgc	gcgtgaacga	tatcacccgg	gacttcacgt	cgatggctcc	cctggacatc	12960
gacaccactc	gcgacaagag	cttcgaacag	cgcgctaagc	gtattcaaga	gcagctgtgg	13020
gaagcgatgg	atcactgcga	cgtaagcggt	atcgaggtcc	agcgagaggc	cgcccgggtc	13080
ctggggatcc	aacgaggcgc	attgttcccc	gtgggtgctca	cgagcgcgct	taaccagcaa	13140
gtcgttggtg	tcacctcggt	gcagaggctc	ggaactccgg	tgtacaccag	cacgcagact	13200
cctcagctgc	tgctggatca	tcagctctac	gagcacgatg	gggacctcgt	cctcgcgtgg	13260
gacatcgctg	acggagtgtt	cccgcgccag	cttctggacg	acatgctcga	agcgtacgtc	13320
gtttttctcc	ggcggtcac	tgaggaaacca	tggggtgaac	aggtgcgctg	ttcgcttccg	13380
cctgcccagc	tagaagcgcg	ggcgagcgca	aacgcgacca	acgcgctgct	gagcgagcat	13440
acgctgcacg	gcctgttcgc	ggcgcgggtc	gagcagctgc	ccatgcagct	cgccgtgggtg	13500
tcggcgcgca	agacgctcac	gtacgaagag	ctttcgcgcc	gttcgcggcg	acttggcgcg	13560
cggctgcgcg	agcagggggc	acgcccgaac	acattggctg	cgggtggatg	ggayaaaggc	13620
tgggagcagg	ttgtcccggt	tctcgcgggtg	ctcgagtcag	gcgcggcccta	cgtyccgac	13680
gatgccgacc	taccggcgga	gcgtatccac	tacctcctcg	atcatggtga	ggtaaagctc	13740
gtgctgacgc	agccatggct	ggatggcaaa	ctgtcatggc	cgccggggat	ccagcggtcg	13800
ctcgtgagcg	aggccggcgt	cgaaggcgag	ggcgaccagc	ctccgatgat	gccccattcag	13860
acaccttcgg	atctcgcgta	tgctcatctac	acctcgggat	ccacaggggt	gccccagggg	13920
gtgatgatcg	atcatcgagg	tgccgtcaac	accatcctgg	acatcaacga	gcgcttcgaa	13980
atagggcccg	gagacagggg	gctggcgctc	tcctcgcgtg	gcttcgatct	ctcggctctat	14040
gatgtgttcg	ggatcctggc	ggcgggcggt	acgatcgtgg	tgccggagcg	gtccaagctg	14100
cgcgatccgg	cgcattgggc	agagttgatc	gaacgagaga	aggtgacggt	gtggaactcg	14160
gtgccggcgc	tgatgcggat	gctcgtcgag	cattttgagg	gtcgcgccga	ttcgctcgct	14220

aggtctctgc	ggctttcgc	gctgagcggc	gactggatcc	cggtgggcct	gcctggcgag	14280
ctccaggcca	tcaggcccgg	cgtgtcgggtg	atcagcctgg	gcggggccac	cgaagcgtcg	14340
atctggtcca	tcgggtaccc	cgtgaggaa	gtcgacctat	cgtgggcgag	catcccctac	14400
ggccgtccgc	tgcgcaacca	gacgttccac	gtgctcgatg	aggcgtcga	accgcgccc	14460
gtctgggttc	cggggcaact	ctacattggc	ggggtcgggc	tggcactggg	ctactggcgc	14520
gatgaagaga	agacgcgcaa	gagcttccctc	gtgcaccccc	agaccgggga	gcgcctctac	14580
aagaccggcg	atctggggccg	ctacctgccc	gatggaaaca	tcgagttcat	ggggcggtgag	14640
gacaacaaaa	tcaagcttcg	cggataccgc	gttgagctcg	gggaaatcga	ggaaacgctc	14700
aagtcgcatc	cgaacgtacg	cgacgcgggtg	attgtgccc	tcgggaacga	cgcggcgaac	14760
aagctccttc	tagcctatgt	ggtcccggag	ggcacacgga	gacgcgctgc	cgagcaggac	14820
gcgagcctca	agaccgagcg	gatcgacgcg	agagcacacg	ccgccgaagc	ggacggcttg	14880
agcgacggcg	agaggggtgca	gttcaagctc	gctcgacacg	gactccggag	ggacctggag	14940
ggaaagcccc	tcgtcgatct	gaccgggcag	gatccgcggg	aggcggggct	ggacgtctac	15000
gcgcgtcgcc	gtagcgtccg	aacgttccct	gaggccccga	ttccggttgt	tgagtttggg	15060
cgattcctga	gctgcttgag	cagcgtggag	cccgacggcg	cgacccttcc	caaattccgt	15120
tatccatcgg	cgggcagcac	gtacctgggtg	caaacctacg	cgtatgtcaa	atccggccgc	15180
atcgagggcg	tggacgaggg	cttctattat	taccaccctg	tcgagcaccg	tttgctgaag	15240
ctctccgatc	acgggatcga	gcgcggagcg	cacgttcggc	aaaacttcga	cgtgttcgat	15300
gaagcggcgt	tcaacctcct	gttcgtgggc	aggatcgacg	ccatcgagtc	gctgtatgga	15360
tcgtcgtcgc	gagaattttg	cctgctggag	gccggatata	tggcgcagct	cctgatggag	15420
caggcgcctt	cctgcaacat	cggcgtctgt	ccggtggggc	aattcaattt	tgaacaggtt	15480
cgcccggttc	tcgacctgcg	acattcggac	gtttacgtgc	acggcatgct	ggcgggcgcg	15540
gtagaccgcg	ggctgttcca	ggctgttagc	ctcggtcagg	attcctcacc	gaggcgcgcc	15600
acgacgcgcg	gcgcctctcc	cgccgcgag	cagcacttcg	ccgatatgct	tcgcgacttc	15660
ttgaggacca	aactaccgga	gtacatgggtg	cctacagtct	tcgtggagct	cgatgcgttg	15720
ccgctgacgt	ccaacggcaa	ggtcgatcgt	aaggccctgc	gcgagcggaa	ggatacctcg	15780
tcgccgcggc	attcggggca	cacggcgcca	cggagccct	tggaggagat	cctcgtcgcg	15840
gtcgtacggg	aggtgctcgg	gctggagggtg	gtcgggctcc	agcagagctt	cgctgatctt	15900
ggtgagacat	cgattcacat	cgttcgcgtg	aggagcctgt	tgcagaagag	gctggatagg	15960
gagatcgcca	tcaccgagtt	gttccagtag	ccgaacctcg	gctcgctggc	gtccgggttg	16020
cgccgagact	cgagagatct	agatcagcgg	ccgaacatgc	aggaccgagt	ggaggttcgg	16080
cgcaagggca	ggagacgtag	ctaagagcgc	cgaacaaaac	caggccgagc	gggcccgatga	16140
gcgcgaagcc	cgcttgcgtc	accctgggac	tcattctgatc	tgatcgcggg	tacgcgtcgc	16200
gggtgtgcgc	gttgagccgt	gttgttccgaa	cgctgaggaa	cggtgagctc	atggaagaac	16260
aagagtccctc	cgctatcgca	gtcatcgcca	tgtcggggcg	ttttccgggg	gcgcggggatc	16320
tggacgaatt	ctggaggaac	cttcgagacg	gcacggaggc	cgtgcagcgc	ttctccgagc	16380
aggagctcgc	ggcgctccgga	gtcgaccccg	cgctgggtgt	ggacccgagc	tacgtccggg	16440
cgggcagcgt	gctggaagac	gtcgaccggt	tcgacgctgc	tttcttcggc	atcagcccgc	16500
gcgaggcaga	gctcatggat	ccgcagcacc	ggatcttcat	ggaatgcgcc	tgggaggcgc	16560
tggagaacgc	cggatacgac	ccgacggctt	acgagggtc	tatcggcggtg	tacgcggcg	16620
ccaacatgag	ctcgtacttg	acgtcgaaac	tcacgcagca	cccagcgatg	atgagggtggc	16680
ccggctgggtt	tcagacgttg	atcggcaacg	acaaggatta	cctcgcgacc	cacgtctcct	16740
acaggctgaa	tctgagaggg	ccgagcatct	ccgttcaaac	tgcttgcctc	acctcgctcg	16800
tggcggttca	cttggcggtg	atgagcctcc	tggaccgcga	gtgcgacatg	gcgctggccg	16860
gcgggattac	cgtccggatc	ccccatcgag	cgggctatgt	atatgctgag	gggggcatct	16920
tctctcccga	cgggcatttg	cgggccttcg	acgccaaggc	gaacggcacg	atcatgggca	16980
acggctcgcg	cgttgtcctc	ctgaagccgc	tggaccgggc	gctctccgat	ggtgatcccc	17040
tccgcgcggt	tatccttggg	tctgccacaa	acaacgacgg	agcgaggaag	atcgggttca	17100
ctgcgcccag	tgagggtggg	caggcgcaag	cgatcatgga	ggcgctggcg	ctggcagggg	17160
tcgaggcccc	gtccatccaa	tacatcgaga	cccacgggac	cggcacgctg	ctcgagacg	17220
ccatcgagac	ggcgggcgctg	cgccgggtgt	tcggtcgcga	cgcttcggcc	cggagggtctt	17280
gcgcgatcgg	ctccgtgaag	accggcatcg	gacacctcga	atcgccgggt	ggcatcgccg	17340
gtttgatcaa	gacggctcttg	gcgctggagc	accggcagct	gccgcccagc	ctgaacttcg	17400
agtctcctaa	cccatcgatc	gatttcgcga	gcagccggtt	ctacgtcaat	acctctctta	17460
aggattggaa	taccggctcg	actccggcgc	gtggccggct	cagctcgctt	gggatcgggc	17520
gcaccaacgc	ccatgtcgtg	ctggaggaag	cgcccgcggc	gaagcttcca	gccgcggcgc	17580
cggcgcgctc	tgccgagctc	ttcgtcgtct	cgcccaagag	cgcagcgggc	ctggatgccg	17640
cggcggcacg	gctacgagat	catctgcagg	cgcaccaggg	gatttcgttg	ggcgagcgtg	17700
ccttcagcct	ggcgacgacg	cgcagcccca	tggagcaccg	gctcgcgatg	cgcgcgccgt	17760
cgcgcgaggc	gttgcgagag	gggtcgcagc	cagcggcgcg	aggccagacc	ccgcggggcg	17820
ccgtgcgtgg	ccgctgctcc	ccaggcaacg	tgccgaaggt	ggtcttcgtc	tttcccggcc	17880
agggctctca	gtgggtcggc	atgggcccgc	agctcctggc	tgaggaaacc	gtcttccacg	17940
cggcgctttc	ggcgtgcgac	cgggccatcc	aggccgaagc	tggttggtcg	ctgctccggg	18000
agctcgccgc	cgacgaaggg	tcctcccagc	tcagcgcgat	cgacgtgggtg	cagccgggtgc	18060
tgttcgccct	cgcggtggca	tttgccggcg	tgtggcggtc	gtgggggtgc	gcqpcgacg	18120
tcgtgatcgg	ccacagcatg	ggcgaggtag	ccgcgcgcga	tgtggccggg	gcgctgtcgc	18180

tcgaggatgc	ggtggcgatc	atctgccggc	gcagccggct	gtccggcgcc	atcagcggtc	18240
agggcgagat	ggcggtgacc	gagctgtcgc	tggccgaggc	cgaggcgggc	ctccgaggct	18300
acgaggatcg	ggtgagcgtg	gccgtgagca	acagcccgcg	ctcgacggtg	ctctcgggcg	18360
agccggcagc	gatcggcgag	gtgctgtcgt	ccctgaacgc	gaaggggggtg	ttctgccgtc	18420
gggtgaaggt	ggatgtcgcc	agccacagcc	cgcaggctga	cccgtctgcg	gaggacctct	18480
tggcagccct	gggcgggctc	cggccgggtg	cggctgcggt	gccgatgcgc	tcgacggtga	18540
cgggcgccat	ggtagcgggc	cgggagctcg	gagcgaatta	ctggatgaac	aacctcaggc	18600
agccagtgcg	cttcgccgag	gtagtccagg	cgcagctcca	aggcgcccac	ggctctgttcg	18660
tggagatgag	cccgcacccg	atcctaacga	cttcggctga	ggagatgcgg	cgcgcggccc	18720
agcggggcgg	cgcagcgggtg	ggctcgtctg	ggcgggggca	ggacgagcgc	ccggcgatgc	18780
tggaggcgct	gggcacgctg	tgggcgcagg	gctaccctgt	accctggggg	cggctgtttc	18840
ccgcgggggg	gcggcgggta	ccgctgccga	cctatccctg	gcagcgcgag	cggctactgga	18900
tcgaagcgcc	ggccaagagc	gcccggggcg	atccgcggcg	cgtgcgtgcg	ggcgggtcacc	18960
cgtcctcggc	tgaatgcag	accctgtcaa	cccagacgag	cacgcggctg	tgggagacga	19020
cgtcggatct	caagcggctg	ccgtggctcg	gcgaccaccg	ggtgcaggga	gcggctcgtgt	19080
ttccggggcg	ggcgtacctg	gagatggcga	tttcgtcggg	ggccgaggct	ttgggcgatg	19140
gcccctttga	gataactgac	gtggtgctcg	ccgaggcgct	ggccttcgcg	ggcgacgcgg	19200
cgggtgttgt	ccaggtgggtg	acgacggagg	agcgtcgggg	gcggctgcag	ttccagatcg	19260
cgaagccggg	gccggggcgt	ggccacgcgt	ccttcggggg	ccacgctcgc	ggcgcggttc	19320
tcagagtggg	gcgcaccgag	gtcccggctg	ggcttacgct	ttccgctgtg	cgcgcgcggc	19380
tcaggccagc	catacccgcc	gcggccacct	acgcggagct	gaccgagatg	gggctgcagt	19440
acggccctgc	cttcacgggg	attgctgagc	tatggcgggg	tgaaggcgag	gcgctggggac	19500
gggtacgcct	gcccgcagcg	gccggctcgg	cagcggagta	tcggttgcat	cctgcgctgc	19560
tggacgcgtg	cttcagatc	gtcggcagcc	tcttcgcccc	cagtggcgag	gcgacgccgt	19620
gggtgcccgt	ggagtggggc	tcgctgcggc	tcttgacgag	gccttcgggg	gagctgtggt	19680
gccatgcgcg	cgtcgtgaac	catggggcag	aaacccccga	tcggcagggc	gccgactttt	19740
gggtgggtcga	cagctcgggt	gcagtggctg	ccgaagtgtg	cgggctcgtg	gcgcagcggc	19800
ttccgggagg	ggtgcgcggg	cgcaagaag	acgattgggt	cctggagctc	gagtggggaa	19860
ccgcagcggg	cggcacagcc	aaggtcaacg	cgggcccggg	gctgctcctc	ggcggcgggc	19920
gtgggctcgg	cgcgcggttg	cgcgcgatgc	tggaggcccg	cggccatgcc	gtcgtgcatg	19980
cggcagagaa	caacacgagc	gctgcggcg	tacgcgcgct	cctggcaaag	gcctttgacg	20040
gccaggctcc	gacggcggtg	gtgcacctcg	gcagcctcga	tgggggtggc	gagctcgacc	20100
cagggctcgg	ggcgcaaggc	gcattggacg	cgcgcgggag	cgcgcagctc	agtcccgatg	20160
ccctcgatcc	ggcgctggta	cgtggctgcg	acagcgtgct	ctggaccgtg	cagggccctgg	20220
ccggcatggg	cttcgagac	gcccgcgat	tgtggctttt	gaccgcgggc	gcacaggccg	20280
tcggcgccgg	cgacgtctcc	gtgacacagg	caccgctgct	ggggctgggc	cgcgtcatcg	20340
ccatggagca	cgcggatctg	cgtgcgctc	gggtcgacct	cgatccagcc	cggcccggagg	20400
gggagctcgc	tgcctgtctg	gccgagctgc	tggccgacga	cgcgaagcg	gaagtgcgct	20460
tgcgcggtgg	cgagcgatgc	gtcgtctcga	tcgtccgcg	gcagcccag	acccggcccc	20520
gggggaggat	cgagagctgc	gttcgcagcc	agctcaccat	ccgcgcggac	agcacctacc	20580
ttgtgaccgg	cggctctgggt	gggctcgggt	tgagcgtggc	cggatggctg	gccgagcgcg	20640
gcgctgggtca	cctgggtgctg	gtgggcccgt	ccggcgccgg	gagcgtggag	caacggggag	20700
ccgtcgcggc	gctcagaggc	cgcggcgccg	gcgtcaccgt	ggcgaaggcg	gatgtcgccg	20760
atcgggcgca	gctcagagcg	atcctccgcg	agttccacc	gtcggggatg	ccgctgcggg	20820
gcgtcgtcca	tgcggccggc	atcttgagcg	acgggctgct	gatgcagcag	actcccgcgc	20880
ggtttcgtaa	ggtgatggcg	cccaaggctc	agggggccct	gcacctgcac	gcgttgacgo	20940
gcgaagcgcc	gctttccttc	ttcgtgctgt	acgcttcggg	agtagggctc	ttgggctcgc	21000
cgggccaggg	caactacgcc	gcggccaaca	cgttctcga	cgctctggcg	caccaccgga	21060
gggcgcaggg	gctgccagcg	ttgagcgtcg	actggggcct	gttcgcggag	gtgggcatgg	21120
cggccgcgca	ggaagatcgc	ggcgcgccgg	tggctccccg	cggaatgcgg	agcctcacc	21180
ccgacgaggg	gctgtccgct	ctggcacggc	tgctcgaag	cggccgcgct	caggtggggg	21240
tgatgccggt	gaacccgcgg	ctgtgggtgg	agctctacc	cgcggcgggc	tcttcgcgaa	21300
tgttgcgcg	cctgggtgag	gcgcacgcg	cgagcgcgg	cgggccaagc	ggggacgggg	21360
acctgctccg	cgcctcgcgc	gctgccgagc	cgagcgcgcg	gagcgcgctc	ctggagccgc	21420
tcctccgcgc	gcagatctcg	caggtgctgc	gcctccccga	gggcaagatc	gaggtggacg	21480
ccccgctcac	gagcctgggc	atgaactcgc	tgatggggct	cgagctgcgc	aaccgcacg	21540
aggccatgct	gggcatacacc	gtaccggcaa	cgtgttgtg	gacctatccc	acggtggcgg	21600
cgctgagcgc	gcatactggc	cgggagggat	gcgaaggcgc	tcctgtggag	tcaccgcaca	21660
ccaccgccga	ctctgcgcgc	gagatcgagg	agatgtcgca	ggacgatctg	acgcagttga	21720
tcgcagcaaa	attcaaggcg	cttacatgac	tactcgcggt	cctacggcac	agcagaatcc	21780
gctgaaacaa	gcggccatca	tcattcagcg	gctggaggag	cggctcgcgt	ggctcgcaca	21840
ggcggagctg	gaacggaccg	agccgatcgc	catcgtcggt	atcggtgcc	gcttccctgg	21900
cgggtgcggc	gctccggaag	cgttttggga	gctgctcgac	gcggagcgcg	acgcgggtcca	21960
gccgctcgac	atgcgctggg	cgtggtggg	tgtcgtctcc	gtcgaggccg	tgccgcactg	22020
ggcggggctg	ctcaccgagc	cgatagattg	cttcgatgct	gcgttcttcg	gcattctcgcc	22080
tcgggaggcg	cgatcgctcg	acccgcagca	tcgtctgttg	ctggaggctg	cttgggaggg	22140



gctcgaggac	gccgggtatcc	cgccccggtc	catcgacggg	agccgcaccg	gtgtgttcgt	22200
cggcgctttc	acggcggact	acgcgcgcac	ggtcgctcgg	ctgccgcgcg	aggagcgaga	22260
cgcgtagacg	gccaccggca	acatgctcag	catcgccgcc	ggacggctgt	cgtacacgct	22320
gggggttgag	ggaccttgcc	tgaccgtcga	cacggcgtgc	tcgtcatcgc	tggtggcgat	22380
tcacctcgcc	tgccgcagcc	tgccgcgcagg	agagagcgat	ctcgcgttgg	cgggaggggt	22440
cagcgcgctc	ctctcccccg	acatgatgga	agccgcggcg	cgcacgcaag	cgctgtcgcc	22500
cgatggctgt	tgccggacct	tcgatgcttc	ggccaacggg	ttcgtccgtg	gcgagggtcg	22560
tggcctggtc	gtcctcaaac	ggctctccga	cgcgcaacgg	gatggcgacc	gcctctgggc	22620
gctgatccgg	ggctcggcca	tcaaccatga	tggccggctc	accgggttga	ccgcgcccaa	22680
cgtgctggct	caggagacgg	tcttgccgca	ggcgctgcgg	agcgcccacg	tcgaagctgg	22740
ggcgcgtgat	tacgtcgaga	cccacggaac	agggacctcg	ctgggcgatc	ccatcgaggt	22800
cgaggcgctg	cggcgacagg	tggggcggcg	gcgctccgac	ggcacacgct	gcgtgctggg	22860
cgcggtgaag	accaacatcg	gccatctcga	ggccgcggca	ggcgtagcgg	gcctgatcaa	22920
ggcagcgctt	tcgctgacgc	acgagcgcac	cccagaaaac	ctcaacttcc	gcacgctcaa	22980
tccgcggatc	cggctcgagg	gcagcgcgct	cgcgttggcg	accgagccgg	tgccgtggcc	23040
gcgcacggac	cgcccgcgct	tcgcgggggt	gagctcgttc	gggatgagcg	gaacgaacgc	23100
gcatgtgggt	ctggaagagg	cgccggcggt	ggagctgtgg	cctgccgcgc	cggagcgctc	23160
ggcggagctt	ttggtgctgt	cgggcaagag	cgagggggcg	ctcgatgcgc	aggcggcgcg	23220
gctgcgcgag	cacctggaca	tgaccccgga	gctcgggctc	ggggacgtgg	cgttcagcct	23280
ggcgacgacg	cgcagtcgga	tgagccacgc	gctcgcgggt	gcggtgacgt	cgcgcgaggg	23340
gctgctggcg	gcgctctcgg	ccgtggcgca	ggggcgagac	ccggcggggg	cggcgcgctg	23400
catcgcgagc	tcctcgcgcg	gcaagctggc	gttcctgttc	accggacagg	gcgcgcagac	23460
gccgggcatg	ggccgggggg	tttgccggcg	gtggccagcg	ttccgggagg	cgttcgaccg	23520
gtgctggggc	ctgttcgacc	gggagctgga	ccgcccgctg	cgcgaggtga	tgtggggcga	23580
ggcggggagc	ggcgagtcgt	tggtgtcgga	ccagacggcg	ttcaccacgc	ccgcgctctt	23640
cgcggtggag	tacgcgctga	cggcgctgtg	gcggtcgtgg	ggcgtagagc	cggagctcct	23700
ggttgggcat	agcatcgggg	agctgggtgg	ggcgctgcgt	gcgggggtgt	tctcgctgga	23760
agatgggggt	aggctcgtgg	cggcgcgcgg	gcggtgatg	caggggctct	cggcggggcg	23820
cgcgatgggt	tcgctcggag	cgccggaggc	ggaggtggcg	gcggcggtgg	cgccgcacgc	23880
ggcgctgggt	tcgatcgcgg	cggtcaatgg	gccggagcag	gtgggtgatc	cgggcggtga	23940
gcaagcgggt	caggcgatcg	cgccgggggt	cgccggcgcg	ggcgcgcgca	ccaagcggct	24000
gcatgtctcg	cacgcgttcc	actcgccgct	gatggaaccg	atgctggagg	agttcggggc	24060
ggtggcgggc	tcggtgacgt	accggcgggc	aagcgtttcg	ctggtgagca	acctgagcgg	24120
gaaggtgggt	acggacgagc	tgagcgcgcc	gggttactgg	gtgcggcacg	tgccggaggc	24180
ggtgcgcttc	gcggacgggg	tgaaggcgct	gcacgaagcc	ggcgcgggga	cgttcgtcga	24240
agtgggcccc	aagccgacgc	tgctcgggct	ggtgccagcc	tgccctgccg	aggcggagcc	24300
gacgctgctg	gcgtcgttgc	gcgcggggcg	cgaggaggct	gcgggggtgc	tcgaggcgct	24360
gggcaggctg	tgggccgcgg	gcggtcgggt	cagctggccg	ggcgtcttcc	ccacggctgg	24420
cgccgggggt	cgcgtcgcca	ccctacgctg	gcagcgccag	cggtactgga	tcgaggcgcc	24480
ggccgaaggg	ctcgagacca	cgcccgccga	tgcgctggcg	cagtggttct	accgggtgga	24540
ctggcccagag	atgcctcgct	catccgtgga	ttcgccggca	gcccggctcc	gcgggtggct	24600
ggtgctggcc	gaccgggggt	gagtcgggga	ggcgcccgcg	gcggcgcttt	cgctgcaggg	24660
atgttcgttg	gccgtgtctc	atgcgcccgc	cgaggctctc	gcgggttgcc	agcagggtac	24720
ccagggccctc	ggtggccgca	gggggtcctg	gcccggctctg	tacctgtggg	gtctggagcg	24780
cgctcgtggag	gcgggggcat	cgcccggaaga	ggtcgccaaa	gtcaccatc	ttgcgcggcg	24840
gocggtgctc	gcgctgatcc	aggcgtcgg	cacggggccg	cgctcaccct	ggctctggat	24900
cgtgacccga	ggggccctgca	cggtggggcg	cgagcctgac	gctgccccct	gtcaggcgcc	24960
gctgtggggg	atgggcccgg	tcgcggcgct	agagcatccc	ggctcctggg	gcgggctcgt	25020
ggacctggat	ccggaggaga	gcccgcagga	ggtcgaggcc	ctggtggccg	agctgctttc	25080
gccggacgcc	gaggatcagc	tggcattccg	ccaggggcgc	cgccgcgcag	cgcggttgt	25140
ggccgccccca	ccggaggga	acgcagcgcc	ggtgtcgctg	tctgcggagg	ggagttactt	25200
ggtgacgggt	gggctggggc	cccttggcct	cctcgttgcg	cgggtggttg	tggagcgccg	25260
ggcggggcac	cttgtgctga	tcagccggca	cggaattgcc	gaccgcgagg	aatggggccg	25320
agatcagccg	ccagaggtgc	gcgcgcgcac	tgccggcgatc	gaggcgctgg	aggcgcaggg	25380
cgccgggggtc	accgtggcgg	cggtcgacgt	ggccgatgcc	gaaggcatgg	cggcgctctt	25440
ggcggccgctc	gagccgcgcg	tgccgggggt	agtgcacgcc	gcgggtctgc	tcgacgacgg	25500
gctgctggcc	caccaggacg	ctggtcggct	cgcccggggt	ttgcgccccca	aggtggaggg	25560
ggcatgggtg	ctgcacaccc	ttaccgcgga	gcagccgctg	gacctcttcg	tactgttttc	25620
ctcgcgctcg	ggcgctcttc	gctcgatcgg	ccagggcagc	tacgcggcag	gcaatgcctt	25680
tttgagcgcg	ctggcggaac	tcgcgcgaac	gcaggggctc	gccgccctga	gcatcgccct	25740
gggcctgtgg	gcggaggggg	ggatgggctc	cagggcgcag	cgccgggaac	acgaggcatc	25800
gggaatcctg	gcgatgccga	cgagtcggcg	cctggcgccg	atggaatggc	tgctcggtac	25860
gcgcgcgacg	cagcgcgtgg	tcattccagat	ggattggggc	catgcgggag	cggcgcggcg	25920
cgacgcgagc	cgaggccgct	tctgggatcg	gctggtaact	gccacgaaag	aggcctcctc	25980
ctcgcccgctg	ccagctgtgg	agcgctggcg	caacgcgctc	gttggtggaga	cccgcctcgg	26040
gctctacag	cttgtgcgcg	gcgtggctcg	cggggtgatg	ggctttaccg	accagggcac	26100

gctcgacgtg	cgacgaggct	tcgccgagca	gggcctcgac	tccctgatgg	ccgtggagat	26160
ccgcaaaccg	cttcagggtg	agctgggtat	gccgctgtcg	gcgacgctag	cgttcgacca	26220
tccgaccgtg	gagcggctgg	tgaataactt	gctgagccag	gcgctggagc	tgaggaccg	26280
caccgacgtg	cggagcgttc	ggttgccggc	gacagaggac	ccgatcgcca	tcgtgggtgc	26340
cgctcgccgc	ttcccgggcg	gggtcgagga	cctggagtcc	tactggcagc	tggtgaccga	26400
ggcggtgggtg	gtcagcaccg	aggtgccggc	cgaccggtgg	aatggggcag	acgggcgcg	26460
ccccggctcg	ggagaggcac	agagacagac	ctacgtgccc	aggggtggct	ttctgcgcca	26520
ggtggagacg	ttcgatgcgg	cgttcttcca	catctcgcc	cgggaggcga	tgagcctgga	26580
cccgaacag	cggctgctgc	tggaaagtga	ctgggaggcg	atcgagcgcg	cgggccagga	26640
cccgctcgccg	ctgcgcgaga	gccccacggg	cggtgttcgtg	ggcgcgggcc	ccaacgaata	26700
tgccgagcgg	gtgcaggaac	tcgccgatga	ggcgcggggg	ctctacagcg	gcaccggcaa	26760
catgctcagc	ggtgcgggcg	gacggctatc	atttttctctg	ggcctgcacg	ggccgacct	26820
ggctgtggat	acggcggtgc	cctcgctcgt	ggtggcgctg	cacctcggt	gccagagctt	26880
gcgacggggc	gagtgcgacc	aagccctggt	tggcggggtc	aacatgctgc	tctcgccgaa	26940
gaccttcgcg	ctgctctcac	ggatgcacgc	actttcgccc	ggcgggcggt	gcaagacgtt	27000
ctcgcccgac	gcgagcggt	acgcgcgggc	cgagggtgc	gccgtgggtg	tgctcaagcg	27060
gctctccgac	gcgagcgcg	accgcgaccc	catctggcg	gtgatccggg	gtacggcgat	27120
caatcatgat	ggcccagaca	gcgggctgac	agtgccagc	ggccctgccc	aggaggcgct	27180
gttacgccag	gcgctggcgc	acgcaggggt	ggttccggcc	gacgtcgatt	tcgtggaatg	27240
ccacgggacc	gggacggcgc	tgggcgaccc	gatcgaggtg	cgtagcgctga	gcgacgtgta	27300
cgggcaagcc	cgccctgccc	accgacccgt	gatacctgga	gccgccaagg	ccaaccttgg	27360
gcacatggag	cccgcggcgg	gcttgccgg	cttgcctcaag	gcggtgctcg	cgctggggca	27420
agagcaaata	ccagcccagc	cggagctggg	cgagctcaac	ccgctcttgc	cgtagggaggc	27480
gctgccgggtg	gcggtggccc	gcgcagcggt	gccgtggccg	cgacgggacc	gcccgcgctt	27540
cgcggggggtg	agctcgttcg	ggatgagcgg	aacgaacgcg	catgtgggtg	tggaaagggc	27600
gccggcggtg	gagctgtggc	ctgccgcgcc	ggagcgtcgc	gcggagcttt	tggtgtgtgc	27660
gggcaagagc	gagggggcgc	tcgatgcgca	ggcgcgcgcg	ctgcgcgagc	acctggacat	27720
gcacccggag	ctcgggctcg	gggacgtggc	gttcagcctg	gcgacgacgc	gcagcgcgat	27780
gaaccaccgg	ctcgcggtgg	cggtgacgtc	gcgcgagggg	ctgctggcgg	cgcttttcgg	27840
cgtagggcgag	gggcagacgc	cgccgggggg	ggcgcgctgc	atcgcgagct	cgtagcgcg	27900
caagctggcg	ttcctgttca	ccggacaggg	cgcgacagcg	ccgggcatgg	gccgggggct	27960
ttgcgcggcg	tggccagcgt	tccgggaggc	gttcgaccgg	tgctggcgcg	tgctcgaccg	28020
ggagctggac	cgcccgtgct	gcgaggtgat	gtgggcgggg	ccggggagcg	ccgagtcggt	28080
gttgctcgac	cagacggcgt	tcaccagcc	cgcgctcttc	acggtggagt	acgcgctgac	28140
ggcgctgtgg	cggtcgtggg	gcgtagagcc	ggagctgggtg	gctgggcata	gcgcccggga	28200
gctggtggcg	gcgtgcgtgg	cgggggtgtt	ctcgctggaa	gatgggggtga	gggtcggtgg	28260
ggcgcgcggg	cggtgatgc	aggggctctc	ggcgggcggc	gcgatggtgt	cgctcgggag	28320
gccggaggcg	gaggtggcgg	cggcgggtgg	gccgcacgcg	gcgtcggtgt	cgatcgcggc	28380
ggtcaatggg	ccggagcagg	tggtgatcgc	ggcgctggag	caagcggtgc	aggcgatcgc	28440
ggcggggttc	gcggcgcgcg	caagcgcgctg	gttcggggcg	catgtctcgc	acgcgtccca	28500
ctcgccgctg	atggaaccga	tgctggagga	ggttcggggcg	gtggcgcgct	cggtgacgta	28560
ccggcgggcca	agcgtttcgc	tggtgagcaa	cctgagcggg	aaggtggtcg	cggacgagct	28620
gagcgcgccg	gggtactggg	tgcggcacgt	gcgggaggcg	gtgcgcttcg	cggacggggt	28680
gaaggcgctg	cacgaagccg	gtgcgggcac	gttcgtcgaa	gtgggcccga	agccgacgct	28740
gctcgggctg	ttgccagcct	gcctggcgga	ggcgagccg	acgctgctgg	cgctcggtgc	28800
cgccggggcgc	gaggaggctg	cgggggtgct	cgaggcgctg	ggcaggctgt	gggcccggcg	28860
cggtcgggtc	agctggccgg	gcgtcttccc	cacggctggg	cggcggtgct	cgctgccgac	28920
ctatccgtgg	cagcggcagc	ggtactggcc	cgacatcgag	cctgacagcc	gtcgccacgc	28980
agccgcggat	ccgacccaag	gctggttcta	tcgctgggac	tggccggaga	tacctcgag	29040
cctccagaaa	tcagaggagg	cgagccgcgg	gagctggctg	gtattggcgg	ataagggtgg	29100
agtcggcgag	gcggtcgctg	cagcgctgtc	gacacgtgga	cttccatgcg	tcgtgctcca	29160
tgccggcgca	gagacatccg	cgaccgccga	gctggtgacc	gaggctgccg	gcggtcgaag	29220
cgattggcag	gtagtgtctt	acctgtgggg	tctggacgcc	gtcgtcggtg	cggaggcgct	29280
gatcgatgag	atcggcgacg	cgaccgctcg	tgctaccgcg	ccggtgctcg	gcttggtcgc	29340
gtttctgagc	accgtgtctt	gttcgccccg	actctgggtc	gtgaccggcg	gggcatgcat	29400
cgttggcgac	gagcctgcga	tcgccccttg	tcaggcgggc	ttatggggca	tgggcccgggt	29460
ggcgggcgctc	gagcatcccc	gggctctggg	cgggctcgtg	gacctggatc	cccagcgag	29520
cccccccaaa	gccagccga	tcgacggcga	gatgctcgtc	accgagctat	tgctcgagga	29580
gaccaggat	cagctcgcc	tccgccatgg	gcgcccgcac	gcggcacggc	tggtggccgc	29640
cccgccacag	gggcaagcgg	caccgggtgtc	gctgtctgcg	gaggcgagct	acctgggtgac	29700
gggaggcctc	ggtgggctgg	gcctgatcgt	ggcccagttg	ctggtggagc	tgggagcgcg	29760
gcacttgggtg	ctgaccagcc	ggcgcggggt	gcccagccgg	caggcggtgt	gcgagcagca	29820
gcccctgag	atccgcgcgc	ggatcgagc	ggctgaggcg	ctggaggcgc	ggggtgcacg	29880
ggtgaccgtg	gcagcggtgg	acgtggccga	cgctgaaccg	atgacagcgc	tggtttcgtc	29940
ggtcgagccc	ccgctgcgag	gggtggtgca	cgccgctggc	gtcagcgctca	tgctccact	30000
ggcgagagcg	gacgagaccc	tgctcgagtc	ggtgctcgt	cccaagggtg	ccgggagctg	30060



gctgctgcac	cggtgctgc	acggccggcc	tctcgacctg	ttcgtgctgt	tctcgtcggg	30120
cgcagcgggtg	tggggtagcc	atagccagg	tgcgtacgcg	gcgcccaacg	ctttcctcga	30180
cgggctcgcg	catcttcggc	gttcgcaatc	gctgcctgcg	ttgagcgtcg	cgtgggggtct	30240
gtggggccgag	ggaggcatgg	cggacgcgga	ggctcatgca	cgtctgagcg	acatcgggggt	30300
tctgcccatg	tgcacgtcgg	cagcgttgct	ggcgctccag	cgcttggtgg	agaccggcg	30360
ggctcagcgc	acgggtgacc	ggatggactg	ggcgcgcttc	gcgccgggtgt	acaccgctcg	30420
agggcgctgc	aacctgcttt	cggcgctgg	cgcaggcg	gacatcatcg	cgccttcccc	30480
tccggcggca	gcaaccggga	actggcggtg	cctgtccggt	gcggaagccc	gctgggtct	30540
gcacgagatc	gtccatgggg	ccgtcgctcg	gggtgctggc	ttcctcgacc	cgagcgcgct	30600
cgatccctggg	atgggggttca	atgagcagg	cctcgactcg	ttgatggcgg	tggagatccg	30660
caacctcctt	caggctgagc	tggacgtgcg	gctttcgacg	acgctggcct	ttgatcatcc	30720
gacgggtacag	cggctgggtg	agcatctgct	cgctcgatgta	ctgaagctgg	aggatcgag	30780
cgacacccag	catgttcggt	cgttggcgtc	agacgagccc	atcgccatcg	tgggagccg	30840
ctgccccttc	ccgggcccggg	tggaggacct	ggagtccctac	tggcagctat	tggccgagg	30900
cgtgggtggc	agcgccgagg	tggccggcga	ccggtgggat	gcgccggact	ggtacgacct	30960
tgatccggag	atcccaggcc	ggacttacgt	gaccaaaggc	gccttccctgc	gcgatttga	31020
gagattggat	gcgaccttct	tccgcatctc	gcctcgcgag	gcgatgagcc	tcgaccgca	31080
gcagcggtt	ctcctggagg	taagctggga	agcgctcgag	agcggggta	tcgctccgga	31140
tacgctgcga	gatagcccca	ccggggtgtt	cgtgggtgcg	gggcccattg	agtactacac	31200
gcagcggtg	cgaggcttca	ccgacggagc	ggcagggttg	tacggcggca	ccgggaacat	31260
gctcagcgtt	acggctggac	ggctgtcggt	tttccctgggt	ctgcacggcc	cgacgctggc	31320
catggatacg	gcgtgctcgt	catccctggt	cgcgtgcac	ctgcctgcc	agagcctgcg	31380
actgggcgag	tgcgatcaag	cgttggtgg	cggggtcaac	gtgctgctcg	cgccggagac	31440
cttcgtgctg	ctctcacgga	tgcgcgcgct	ttcgcccagc	ggcggtgca	agacgttctc	31500
ggccgacgcg	gacggctacg	cgccgggcga	ggggtgcg	gtggtggtgc	tcaagcggct	31560
gcgcgatgcg	cagcgcccg	gcgactccat	cctggcgctg	atccgggaa	gcgcggtgaa	31620
ccacgacggc	ccgagcagcg	ggctgacctg	acccaacgga	cccggccagc	aagcattgct	31680
gcgccaggcg	ctttcgcaag	caggcgtgtc	tccggtcgac	gttgattttg	tggagtgtca	31740
cgggacagg	acggcgctgg	gcgacctgat	cgaggtgcag	gcgctgagcg	aggtgtatgg	31800
tccaggggcg	tccggggacc	gaccgctggt	gctggggggc	gccaaggcca	acgtcgcgca	31860
tctggaggcg	gcacttggtc	tggccagcct	gtgcttgccg	gtgcttgccg	tgccgacga	31920
gcagatcccc	gcccagcccg	agctggggga	gctcaaccg	cacttgccgt	ggaacacgct	31980
gcccgtggcg	gtgccacgta	aggcggtgcc	gtggggggcg	ggcgacgccc	cgcgctcgggc	32040
cggcgtagcg	gcgttcgggt	tgagcggaac	caacgtgcat	gtcgtgctgg	aggaggcacc	32100
ggaggtggag	ccggcgcccg	cggcgccggc	gcgacgggtg	gagctggctg	tgctatcgcc	32160
caagagcgcg	gcggcgctgg	acgcgcggcg	ggcagggctc	tcggcgccac	tgtccgcgca	32220
cccggagctg	agcctcgggc	acgtggcggt	cagcctggcg	acgacgcgca	gcccgatgga	32280
gcaccggctc	gccatcgcg	cgacctcgcg	cgaggccctg	cgaggcgcg	tggacgcgcg	32340
ggcgagcaa	aagacgcgcg	agggcgcggt	gcgcggaag	gcccgtgtct	cacgcggtaa	32400
gctggctttc	ctgttcaccg	gacaggcg	gcaaattgcc	ggcatggggc	gtgggctgta	32460
cgaaacgtgg	cctgcgttcc	gggaggcggt	cgaccgggtg	gtggcgctct	tcgatcgggg	32520
gatcgaccag	cctctgcg	aggtgatgtg	ggctgcgcg	ggcctcgctc	aggcgcgcg	32580
gctcgatcag	accgcgtacg	cgcagccggc	tctctttg	ctggagtacg	cgctggctgc	32640
cctgtggcgt	tgcgtggggc	tggagccgca	cgctactgct	ggtcatagca	tcggcgagct	32700
ggctgcggcc	tgcgtggcg	gcgtgttctc	gcctgaagat	gcggtaggt	tgggtggccg	32760
gcgcggggcg	ctgatgcagg	cgctaccgcg	cggcggtgcc	atggtagcca	tcgcagcgct	32820
cgaggccgag	gtggccgcct	ccgtggcgcc	ccacgcggcc	acgggtgtcg	tcgcccgggt	32880
caacggctct	gacgccgtcg	tgatcgccgg	cgccgaggta	caggtgctcg	ccctcgccg	32940
gacgttcg	gcgcgtggga	tacgcacgaa	gaggtcgcc	gtctcccatg	cgctccactc	33000
gccgctcatg	gatccgatgc	tggaaagact	ccagcggtc	gctgcgacga	tcgctgaccg	33060
cgccccagac	cgcccggtgg	tgtcgaatgt	caccggccac	gtcgaggcc	ccgagatcgc	33120
cacgcccag	tattgggtcc	ggcatgtgcg	aagcgccgtg	cgcttcggcg	acggggcaaa	33180
ggcgttgc	gcccgggtg	ccgccacggt	cgctgaggtt	ggcccgaagc	cggtcctgct	33240
cgggctgtg	ccagcggtg	tcggggagc	ggcagcggtc	ctcgtgccgt	cgctacgcgc	33300
ggaccgctcg	gaatgcgagg	tggctcctcg	ggcgctcggg	gcttggtatg	cctggggggg	33360
tgcgctcgac	tggaaaggcg	tgttccccga	tggcgcgcg	cgcggtggctc	tgcccattgta	33420
tccatggcag	cgtgagcgcc	attggatgga	cctcaccg	cgaagcgccg	cgccgtgag	33480
gatcgaggt	cgctggccgc	tggctggtgt	cgggctctgc	atgcccggcg	ctgtgttgca	33540
ccacgtgctc	tcgatcgga	cacgccatca	gcccttctc	ggtgatcacc	tcgtgtttgg	33600
caagtggtg	gtgcccggcg	cctttcatgt	cgcggtgatc	ctcagcatcg	ccgcccagcg	33660
ctggcccag	cgggcatcg	agctgacagg	cgtggagttc	ctgaaggcca	tcgcatgga	33720
gcccagacc	gaggtcgagc	tccacgcccgt	gctcaccg	gaagccggcg	gggatggcta	33780
cctgttcgag	ctggcgaccc	tggcgcgcc	ggagaccgaa	cgccgatgga	cgaccacgc	33840
ccgcggtcg	gtgcagccga	gcccggcgcg	ttgcccgcg	tcgaggtgct	tcgaggtgct	33900
ggaggaccgc	gcgatccagc	ccctcgactt	cgccggattc	ctcgacaggt	tatcgcggt	33960
gcggtacggc	tggggtccgc	tttggcgatg	gctgcaggac	ggcgcgctcg	gcgacgaggc	34020

ctcgtttgcc	accctcgtgc	cgacctatcc	gaacgcccac	gacgtggcgc	ccttgcaccc	34080
gacctgtctg	gacaacggct	ttgcggtgag	cctgctgtca	acccggagcg	agccggagga	34140
cgacgggacg	cccccgctgc	cgttcgccgt	ggaacgggtg	cggtggtggc	ggcgccgggt	34200
tggaaggggtg	cggtgtggcg	gcgtgccgcg	gtcgcaggca	ttcgggtgtct	cgagcttcgt	34260
gctggtcgac	gaaactggcg	aggtggtcgc	cgaggtggag	ggatttgttt	gccgccgggc	34320
gccgcgagag	gtgttcctgc	ggcaggagtc	ggcgcgctcg	actgcagcct	tgtaccgct	34380
cgactggccc	gaagcgccct	tgcccgatgc	gcctgcggaa	cggatcgagg	agagctgggt	34440
cggtggtggca	gcacctggct	cggagatggc	cgcgccgctc	gcaacacggc	tcaaccgctg	34500
cgtcctcgcc	gaacccaaag	gcctcgaggc	ggccctcgcg	ggggtgtctc	ccgcaggtgt	34560
gatctgcctc	tgggaggctg	gagcccacga	ggaagctccg	gcgccggcgc	agcgtgtggc	34620
gaccgagggc	ctctcggtgg	tgcaggcgct	cagggaccgc	gcggtgcgcc	tgtggtgggt	34680
gacctggggc	gcagtggccg	tcgaggccgg	tgagcgggtg	caggtcgcca	cagcgccgggt	34740
atggggcctc	ggccggacag	tgatgcagga	gcgcccgag	ctcagctgca	ctctggtgga	34800
tttggagccg	gaggccgatg	cagcgcgctc	agctgacgtt	ctgttgccgg	agctcggtcg	34860
cgctgacgac	gagacacagg	tggctttccg	ttccggaaag	cgccgcgtag	cgcggtgtgt	34920
caaagcgacg	acccccgaag	ggctcctggt	cctgacgca	gagtcctatc	gactggaggc	34980
tgggcagaag	ggcacattgg	accagctccg	cctcgcgccg	gcacagcgcc	gggcacctgg	35040
cccgggcgag	gtcgagatca	aggtaaccgc	ctcggggctc	aacttccgga	ccgtcctcgc	35100
tgtgctggga	atgtatccgg	gcgacgcggg	gccgatgggc	ggagattgtg	ccggtgtcgc	35160
cacggcggtg	ggccaggggg	tgcgccacgt	cgcggctggc	gatgtgttca	tgacgttggg	35220
gactcccgcg	cgattcgtca	cggctgcagc	gcggctgggt	gtccggcagc	ctgcagggt	35280
cgacctgggg	aatctgcggc	gcggcgagcg	cgcttctctg	acggcctggc	tcgtcttgca	35340
gggcatggcc	gcggtgcaaa	tcgcccgatg	ggtgctgac	catgctgcgg	ccggcggtgt	35400
cccgtccaag	tgggcagcgg	ttcaggccat	gataggggcc	gaggtgttcg	ccacggcgag	35460
gcggacgctg	gagtttgctg	agacgttccg	gggctgcgg	cgacgcaca	tcgccagctc	35520
ggtgctcaac	gcgtggcccg	gcgagttcgt	ggacgcgagc	ggcgcccggg	gcgtggacgt	35580
cgggcggttc	ctcgagatgg	gcaagaccga	catacgggat	ctgtccctgc	tgtcgacggg	35640
gcatcccggt	gttcgctatc	gggtattcga	catcctggag	cgagcccgcg	tcgcggcgcc	35700
agagatccct	gagcgcgtgg	tcgagggtct	tgctgcggga	ctcgctccgg	atcgaactcg	35760
gcatgcgttc	gcgatcacca	aggccctcgg	agccttccgg	catctgcgcg	cattgccggg	35820
tcaggggcaag	gtcgtgctgc	tgccggcgcc	ctccgcagcg	ttcatggcgc	aagcgcgcca	35880
cgtactgctg	accggtgggc	tgggagcggt	ggggctccac	cccttggcgc	cgacggggcac	35940
gcagggcggtg	ccgcacatgg	tgctcacagg	tcggcggggc	gtggcccgcg	ggctcgccca	36000
caaagccgct	gcggagatcg	aagcgctcgg	cgcctcgggtg	ctggatacgc	cgggcgctgc	36060
cgccgatcgg	aatgcgctgg	aggctgtgct	ccaggccatt	acgatcgcgg	cgctcggtgt	36120
gggcgtgatc	catgcagccg	gagcgctcga	tgatggtgtg	ccggcgaggt	ggccgttaca	36180
ccgcttctcg	cgggtgctgg	caccgaagggt	gactggcgcc	cttgatgagc	agaccaccga	36240
ggcgggcaac	gatctcgctt	tcttcgtgct	gttctcctcc	tggaaatctgc	atgagctcac	36300
ggccgggcag	tcacaactatg	cggcgggccaa	gaccttctc	atgtcggggc	tcttgggctc	36360
gcgggcccga	ggcctggcgg	cgagagccct	cgctgggggc	gacgcgctgg	ccgcgcacgc	36420
ggcagcgggg	ctcagcgcg	cgctgcaggc	gcggctcgct	ccatggtcgg	acggaggcat	36480
gtcgccggct	cagggcaccg	cgctgctcgg	gcaggcgctg	cgccatggga	tgggagctct	36540
cggggcgatg	tcgctcgacg	tgctgcggc	aagccaagct	gctcgcccg	aaacgcagct	36600
tgtgtggcgc	gcgttggtgc	gcgcggaggc	gcgccatacg	tcgggagcgg	cagtgcggcc	36660
attggccgcg	cgtcttgggg	cgctgcccga	ggcgcgctgc	gcggctgggg	cgaggggggc	36720
cgtgcaggcc	gagatcgcg	gcgtgctttc	atggagcgcc	gccgacgagg	tgcgcaagggt	36780
tcggccgctg	tcggacttgg	gcctcgactc	gctcacggcg	gcgagcgccg	tgcccgtcga	36840
cgccagcgcg	gtgggtgcga	cgctgcgggc	gacgtggcca	gtggagctgc	gcaacgtgct	36900
cgcgctcacg	cgctggctgc	tcgataagggt	cctggccgtg	ttcgatcacc	cgacggtcga	36960
cgcaaagtcg	tcgcccgagg	tcgccctcga	cgagcccat	gccgagccga	gcgtatcgte	37020
ccgtttccca	ggcggcgtgg	ccgatccgga	gtcgttttgg	gccatcatcg	gcatcggtcg	37080
tcgatccgctc	gtcgaggtgc	cgcatgagcg	atgggacatc	cggctgctcg	aagagggcag	37140
tcgggatgtg	cgcggaaga	tgacgacacg	ctttggcgcc	gacgcgttct	atgatccgga	37200
gttcgatccg	gccttcttcg	gcattctgcc	gcgcgaagcg	ttcctgtccg	atatcgaccg	37260
gcggctgctc	ctggagacga	gctgggaggc	gttcgagcgc	acgaccatgg	atccgcagca	37320
gctgatgggc	agcgataccg	gcgtgttcgt	ggggctcttc	gccgggattt	tgcccagagc	37380
cgccggcgcc	atcgaggcgt	tcgatggcta	tctaggcacc	taccaggagt	acgctgcgct	37440
ctcgggcagg	atctcttatg	tgctcgggct	aaaggggccc	ggcaccacgg	ccagcgctgc	37500
gtgctcctcg	tcgctggctg	cggtgcacct	ggcctgccag	agcctgacgg	tggacaccgc	37560
ttcggtggcg	ctggccggcg	gcgtggcgct	gatgtcaccg	gcgctgcggc	ggggcgagtg	37620
cagccggctg	cgaggcctgg	ctcccgcagg	acggtgcgaag	ccggcgacgt	tcgtggagtt	37680
cgcggtgggg	tggagcgaag	gctgcccatt	gctcctgctc	agcttctcgg	ccgcagccga	37740
gcgcgatggg	gatccgatcc	tggcggtgat	ccgcggcacc	aaaccgcttc	gcgatgcgca	37800
cagcaacggg	ctgacggcgc	ccaacgggtc	gtcgcagcaa	gcggtgaacc	aggatggggc	37860
ggagcaggcg	gggctggctc	cgccggacgt	cagctacgtc	gaggtgatcc	gtcggggcct	37920
				gagtgccacg	gcaccggcac	37980

gacgttgggc gaccccatcg aagtgcaggc cctggggcgcc gtgctggcac aggggagacc 38040  
 ctcggaccgg ccgctcgtga tcgggtcggg gaagtccaat atcgacata cgcaggctgc 38100  
 ggcggggcgtg gccggtgtca tcaaggtggc gctggcgctc gagcgcgggc ttatcccgag 38160  
 gagcctgcat ttcgacgcgc ccaatccgca cattccgtgg tcggagctcg ccgtgcaggt 38220  
 ggccgccaaa ccgctcgaat ggacgagaaa cggcgtgccg cgacgagccg gggtagactc 38280  
 gtttggcgte agcgggacca acgcgcacgt ggtgctggag gaggcgccag cggcgcggtt 38340  
 cgcgcccgcg gcggcgcggt cagcggagct tttcgtgctg tcggcgaaga gcgcccggc 38400  
 gctggagcgc caggcgggcg ggctttcggc gcacgtcgtt gcgcacccgg agctcggcct 38460  
 cggcgacctg gcgttcagcc tggcgacgac ccgcagcccc atgacgtacc ggctcgcggt 38520  
 ggcggcgacc tcgcgcgagg cgctgtctgc cgcgctcgac acagcggcgc aggggagagg 38580  
 gccgcccga gccgctcgcg gccacgcttc cacaggcagc gccccaaagg tggttttcgt 38640  
 ctttcctggc cagggtctcc agtggtcggg catgggcaa aagctcctct cggaggagcc 38700  
 cgtcttcggc gacgcgtctt cggcgtgtga ccagcgatt caggccgaag ccggtcgggt 38760  
 gctgctcgcc gagctcggg ccgatgagac cacctcgcag ctgcggccga tcgacgtggt 38820  
 gcagccggcg ctgttcgcga tcgaggtcgc gctgtcggcg ctgtggcggt cgtggggcgt 38880  
 cgagccggat gcagtggtag gccacagcat gggcgaaagt gcggccgcgc acgtcgcgg 38940  
 cgccctgtcg ctcgaggatg ctgtagcgtg catctgcggc cgcagcctgc tgctcggcg 39000  
 gatcagcgcc caaggcgaga tggcggtcgt cgagctttcc ctggccgagg ccgaggcagc 39060  
 gctcctgggc tacgaagacc ggctcagcgt ggcggtgagc aacagcccgc gctcgacggt 39120  
 gctggcgggc gagccggcag cgctcgcaga ggtgctggcg atccttgccg caaaggggg 39180  
 gttctgccgt cgagtcaagg tggacgtcgc cagccacagc ccacagatcg acccgctgcg 39240  
 cgacgagcta ttggcagcat tggcgagct cgagccgcga caagcgacc tgctgatgcg 39300  
 ctcgacggtg acgagcacga tcatggcggg ccggagctc gtggcgagct actgggcgga 39360  
 caacgttcga cagccggtgc gcttcgccga agcggtgcaa tcgttgatgg aagacggtca 39420  
 tgggctgttc gtggagatga gcccgcatcc gatcctgacg acatcggtcg aggagatccg 39480  
 acgggcgacg aagcgggagg gagtcgcggt gggctcgttg cggcgtagac aggcagagcg 39540  
 cctgtccatg ttggaggcgc tgggagcgt ctgggtacac ggccaggcgg tgggctggga 39600  
 gcggctgttc tccgcgggcg gcgcgggct ccgtcgcgtg ccgtgcccga cctatccctg 39660  
 gcagcgcgag cggtagtggg tcgatgcgcc gaccggcggc gcggcgggcg gcagccgctt 39720  
 tgctcatgcg ggcagtcacc cgctcctggg tgaatgcag accctgtcga ccagaggag 39780  
 cacgcgcgtg tgggagacga cgctggatct caaacggctg ccgtggctcg gcgatcaccg 39840  
 ggtgcagggg gcggtcgtgt tcccgggcgc ggcgtacctg gagatggcgc tttcgtccgg 39900  
 ggccgaggcc ttgggtgacg gtccgctcca ggtcagcgat gtggtgctcg ccgaggcgct 39960  
 ggccctcgcg gatgatacgc cggcgcggt gcaggtcatg gcgaccgagg agggaccagg 40020  
 ccgctgcaa ttccacgttg cgagccgggt gccgggcccac ggcggtgctg cctttcgagg 40080  
 ccatgccgc gcggtgctgc gccagatcga cgcgcgcgag gtcccggcga ggctggatct 40140  
 ggccgcgctt cgtgcccggc ttacggccag cgcacccgct gcggctacct atgcggcgct 40200  
 ggccgagatg gggctcagat acggcccagc gttccagggg cttgtcgagc tgtggcgggg 40260  
 ggagggcgag gcgctgggac gtgtcgggt cttccagctg gccggctccc cagcccgctg 40320  
 ccggctccac tggatgcgtt taccgtgga aatcggtcgc ctgcgggtgt tccagcgcc 40440  
 cggcgaggcg acgcatggg atgcgcggag tgtgagccac ggaagccaa caccgagccg 40500  
 gtcgggggag ctgtggtgtc atgcgcggag cagggcgcg atcgtcgccg agatctccg 40560  
 gcggagtacc gacttctggg cgggaggtg acgcccgcgc gaagaagacg actggttcat 40620  
 gctcgtggcg cagcggctcg cgggaggtg ccgctgctcc cggatccgag gtcatggcg 40680  
 ggagccggct tgggaaccga ccgctgctcc tgcgctccac tcggcgctga cggaagctg 40740  
 gctcatcggc tcgggcggcg ggctcggcgc cagagcgcc gccgggttgc aggcactctt 40800  
 ccattccgct ttcgacggc agggcccagc gtcggtggtg cacctcggca gcctcgatga 40860  
 gacggcgctc ctcgacgcgc atgcccctt cgacgccgat gcgcttgagg agtcgctggt 40920  
 gcgtggcggt ctcgacgcgc tctggaccgt gcaggccgtg gccggggcg gcttccgaga 40980  
 gcgcggctgc ttgtggctcg tgacacgcgg cgctcaggcc atcggcgccg gcgacgtctc 41040  
 tcttcgcggg ggcgcgtcc tggggtcgg gcggttacc ccgttgagg gccttgagg acgcccagct 41100  
 tgtggcgcaa gcgctgacc cggatcgacc tcgatccagc gcggcgcgac ggagaagtcg atgagctgct 41160  
 ggcgtgctg ttggccgacg acgcccagga ggaagtgcgc tttcgcggcg gtgagcgcg 41220  
 cgtggcccg ctcgtccgaa ggctgcccga gaccgactgc cgagagaaaa tcgagcccgc 41280  
 ggaaggccgg ccgttccgg cggagatcga tgggtccggc gtgctcgacg acctggtgct 41340  
 ccgagccacg gagcggcgcc ctccctggccc gggcgaggtc gagatcgccg tcgagggcg 41400  
 ggggctcaac tttctcgacc tgatgaggc catggggatc taccctgggc ccggggacgg 41460  
 tccggttgcg ctgggcggcg agtgctccgg ccgaattgtc gcgatggcg aaggtgtcga 41520  
 gagccttcgt atcgccagg acgtcgtggc cgctcgcgcc ttcagtttcg gcacccacgt 41580  
 caccatcgac gcccggtatg tcgcaacctg ccccgcgggc ctgacggcg cggaggcagc 41640  
 cgcgtgccc gtcgcattca tgacggcctg gtacggtctc gtcacatctg ggaggctccg 41700  
 ggccggcgag cgcgtgctca tccactcggc gacggggggc accgggctcg ctgctgtgca 41760  
 gatcggccgc cactcggcg cggagatatt tgcgaccgct ggtacaccgg agaagcgggc 41820  
 gtggctgcgc gagcagggga tcgcgcacgt gatggactcg cggtcgctgg acttcgccga 41880  
 gcaagtgtcg gccgcgacga agggcgaggg ggtcgacgtc gtgttgaaact cgctgtctg 41940

cgccgcgac gacgcgagcc tttcgaccct cgtgccggac ggccgcttca tcgagctcgg 42000  
 caagacggac atctatgcag atcgctcgct ggggctcgct cacttcagga agagcctgtc 42060  
 ctacagcgcc gtcgatcttg cgggcttggc cgtgcgtcgg cccgagcgcg tcgacgct 42120  
 gctggcggag gtgggtggacc tgctcgacg gggagcgctg cagccgcttc cggtagagat 42180  
 cttccccctc tcgcgggccg cggacgcgtt ccggaatg gcgcaagcgc agcatctcgg 42240  
 gaagctcgtg ctcgcgctgg aggaccggga cgtgcggatc cgcgttcggg gcgaatccgg 42300  
 cgctcgccatc cgcgcggagc gcgcctacct cgtgaccggc ggtctggggg ggctcggtct 42360  
 gagcgtggct ggatggctgg ccgagcaggg cctggggcat ctgggtgctg tgggcccgtc 42420  
 cggcgcggtg agcgcgagc agcagacggc tgctcgccg ctcgagggcg acggcgcgcg 42480  
 tgtcacggta gcgagggcag acgtcgccga tcgggcgag atggagcgga tcctcccgca 42540  
 ggttaccgcg tcggggatgc cgctccgcg cgtcgttcat gcggccggaa tcctggacga 42600  
 cgggctgctg atgcagcaaa cccccgcg cgtccgcg cgcgtggcgc ccaaggctcg 42660  
 aggggecttg cacctgcatt cgttgacag tgggctcgcc gggccagggc aactacggcg cggccaacac 42780  
 cgcttcggga gcaaggctct ggggctcgcc gggcagggg ctgccagcat tgagcatcga 42840  
 gttcctcgac gcactggcac accaccggag gggcagggg caaaatcgcg gcgacggct 42900  
 ctggggcctg ttgcgggacg tgggtttggc cgccgggag cgtggggcgc tcgagcgct 42960  
 ggtcaccgct gggacgcgga gcctcacccc catgcccgtt gacgtgcggc agtgggtgga 43020  
 gctcgacggc gatcgacccc cttcgcgag gttgtcgcg ctcattgacg cagggcgct 43080  
 gttctacccg gcggcgcat cttcgcgag gttgtcgcg ctcattgacg cagggcgct 43140  
 ggcttcgggt cggctcgccg gggatcgga cctgctcgaa cggctcgcca cggcgaggc 43200  
 gggcgcgcg gcaggatgc tgcaggaggt cgtgcgcg caggctcgc aggtgctgcg 43260  
 cctctccgaa ggcaagctc acgtggatgc gccgtcacg agcctgggaa tggactcgct 43320  
 gatggggcta gagctgcga accgcacga ggccgtgct gccatcacca tgccggcgac 43380  
 cctgctgtgg acctacccca cgggtggcag gctgagtgcg catctggct ctcatgtcgt 43440  
 ctctacgggg gatggggaat ccgcgcgcc gccgataca gggagcgagg ctccaacgac 43500  
 ccacgaagtc gcttcgctcg acgaagacgg ttgttcgcg accgagaagg ccagctcctg gagcgcttgc 43560  
 gcgcgcggga aagaggtgat tgcgtgacag accgagaagg ccagctcctg gagctcgaga 43620  
 gtgaggttac tctggccctt cgcaagacgc tgaacgagcg cgataccctg gagctcgaga 43680  
 agaccgagcc gatcgccatc gtggggatcg gctgcccgtt ccccgccgga gcggggcactc 43740  
 cggaggcgtt ctgggagctg ctgcacgacg gccgcgacg gatccggccg ctcgaggagc 43800  
 gctgggctg cgtaggtgtc gacccaggcg acgacgtacc gcgctggcg gggctgtcga 43860  
 ccgaggccat cgacggcttc gacgcccgtt tcttcggtat cgcccccg gaggcacgg 43920  
 cgctcgaccc gcagcatcgc ctgctgctgg aggtcgctg ggaggggttc gaagacggc 43980  
 gcatcccgcc caggtccctc gtcgggagcc gcaccgctg gttcgtcgcc gtctgcgcca 44040  
 ccaggagcct ccacgcccgc atcgccgccc gacggctatc gtacacgctg gggctgcagg 44100  
 ccaccggcaa gatcgctcga gagagcgatc tcgctgggc gggaggggtc aacatgcttc 44220  
 gatcggccat attgagcgac gctctgggc gacccaggc gctgtcgccc aatggccgtt 44280  
 agggggcgct cttgcgcgag gctctgggc gacccaggc gggaggggtc ggtctgatcg 44340  
 acatcgagac cgacgcgtcg gcgcggcggg atggggaccg gatctgggcg ctgatccgag 44400  
 gcgctgtggt ggggcccggc gcgctgcgga acgcccggc catcgagatc cgtgctgggc 44460  
 ccaacctcgg ccacctggag ggcgctgccc tgggcagccc cgtgctgggc cctgatcaag 44520  
 cgctacatca cgagcgcatc ccgaggaacc tcaactttcg tacgtcaat ccgcgatcc 44580  
 ggatcgaggg gaccgcgtc gcgttggcga ccgaaccggg gccctggccg cggacgggct 44640  
 ggacgcgctt cgcgggagtg agctcgttcg ggtagagcg gaccaacgcg catgtggtgt 44700  
 tggaggaggc gccggcggtg gagcctgagg ccgcccgc cagagcgcg catgtggtgt 44760  
 tcgtcctgtc ggcgaagagc gcggcgccgc tggatgcgca ggcagcccgg ctgcccggac 44820  
 acctggagaa gcacgtcgag cttggccctg gcgagtggtg gttcagcctg gcgacgacg 44880  
 gcagcgcgat ggagcaccgg ctggcggttg ccgcgagctc cgcgagggcg ctcgagggg 44940  
 cgcttccggc cgcagcgag gtcttcgtgt tccccggta gggctcgag tgggtgggca 45000  
 gggccagcgc gccgaagggt gctcatggcc gaagagccgg tcttccggg ggcgctggag 45060  
 tgggcccga gctcatggcc ggcggaagcg ggctggctgc tgctcgggga gctctccgc 45120  
 gggccatcga ggcggaagcg ggctggctgc tgctcgggga gctctccgc 45180  
 cctcgagct cgggcgcatc gacgtgggtt agccggtgct agccggaagc ggtggtggg 45240  
 tttctcgct gtggcggtcg tggggagtg agccggaagc ggtggtggg 45300  
 gcgaggttgc ggcggcgac gtggccggcg cgtgtcgct cgaggacgcg gtggcgatca 45360  
 tctgcccggc cagccggctg ctgcccggga tcagcggtca gggggagatg gcgctggtc 45420  
 agctgtcgct ggaggaggcc gagggcgcg tcggtggcca tgagggtcgg ctgagcgtg 45480  
 cggtgagcaa cagcccgcg tgcaccgtgc aagggggtgt tctggcgga ggtgaagg 45540  
 tgctggcgcc gctgacggcc ccgctgcgcg aagagctgat cgcggcgctg ggagcgatc 45600  
 gccatagccc gcaggtcgac ccgctgcgct cgacggtgac gggcggggtg atcggggtc 45660  
 ggccgcgagc ggctgcggtg ccgctgcgct cgacggtgac gggcggggtg atcggggtc 45720  
 gggcggggtg atcggggtc cgcggcgctg gggcggggtg atcggggtc 45780  
 gggcggggtg atcggggtc cgcggcgctg gggcggggtg atcggggtc 45840  
 gggcggggtg atcggggtc cgcggcgctg gggcggggtg atcggggtc 45900

cggagctcgg tgcgagctac tgggcgagaca accttcggca gccggtgcgc ttcgctgcgg 45960  
 cggcgcaagc gctgctggag ggtggccccg cgctgttcat cgagatgagc ccgcaccgca 46020  
 tcctgggtgcc gccctgggac gagatccaga cggcgggcca gcaagggggc gctgcggtgg 46080  
 gctcgtcgcg gcgagggcag gacgagcgcg cgacgctgct ggagggcgctg gggacgctgt 46140  
 gggcgctccg ctatccgggtg agctgggctc ggctgttccc cgcgggcggc aggcgggttc 46200  
 cgctgccgac ctatccctgg cagcacgagc ggtgctggat cgaggtcgag cctgacgccc 46260  
 gccgcctcgc cgcagccgac ccaccaagg actggttcta ccgaacggac tggcccggag 46320  
 tgccccgcgc cgccccgaaa tcggagacag ctcatgggag ctggctgctg ttggccgaca 46380  
 ggggtggggg cggtagggcg gtcgctgcag cgctgtcgac gcgcggactt tcctgcaccg 46440  
 tgcttcatgc gtcggtgac gcctccaccg tcgcccagca ggtatccgaa gctgccagtc 46500  
 gccgaaacga ctggcagggg gtcctctacc tgtggggcct cgacgccgtc gtcgatgctg 46560  
 gggcgctcgc cgacgaagtc agcgaggcta cccgccgtgc caccgcaccc gtccttggg 46620  
 tgggtcgatt cctgagcgct gcgccccatc ctctcgtctt ctgggtgggt acccgcgggg 46680  
 catgcacggt gggcgcgag ccagaggcct ctctttgcca agcgcgcttg tggggcctcg 46740  
 cgcgctcgc ggcgctggag caccgcgctg cctgggggtg cctcgtggac ctggatcctc 46800  
 agaagagccc gacggagatc gagcccctgg tggccgagct gctttcgccg gacgccgag 46860  
 atcaactggc gttccgcagc ggtcgcaagg acgcagcacg ccttgtagcc gcccgccgg 46920  
 agggcgacgt cgcaccgata tcgctgtccg cggaggggag ctacctggtg acggcgggc 46980  
 tgggtggcct tggctgtctc gtggctcggt ggctgggtgga gcggggagct cgacatctgg 47040  
 tgctcaccag ccggcacggg ctgccagagc gacaggcgctc gggcgagag cagccgcccg 47100  
 agggccgcgc gcgcatcgca gcggtcgagg gctggaagc gcaggcgcg ggggtgaccg 47160  
 tggcagcggt ggatgtcgcc gaggccagtc ccatgacggc gctgctggcc gccatcgagc 47220  
 cccggttgcg cgggggtggt cagcccgccg gcgtcttccc cgtgcgtcac ctggcggaga 47280  
 cggacgaggg cctgctggag tcggtgtccc gtcccaagg tgggtgctgc 47340  
 accggctgct gcgcgaccg cctctcgacc tgttctgctt gttctcgtcg ggcgcgggc 47400  
 tgtggggtgg caaaggccaa ggcgcatacg cgcggccaa tgcgttcttc gacgggtcg 47460  
 gcaccatcg ccgcgccac tcgctgcccg cgttagacct cgctggggc ttatggggcg 47520  
 agggaggcat ggttgatgca aaggctcatg cacgtctgag cgacatcggg gtcctgccc 47580  
 tggccacggg gccggccttg tcggcgctgg agcgccctgg gaacaccagc gctgtccagc 47640  
 gttcggtcac acggatggac tgggcgcgct cgcgcgggtg ctatgccg cgaggcgccg 47700  
 ggaacttgc ttcggctctg gtgcggagg gtcgctccc tgcgtctccc ccggtgccga 47760  
 cggcaaacg gatctggcg catcgctgcc cgggtgtccg ttgaggagag ccgctcagcc 47820  
 tcgttcgchg gctgctgcc cgggtgtcgg gcttctccga cccggcgcg ctcgacgtcg 47880  
 gccgaggctt cgccgagcag gggctcgact ccctgatggc tctggagatc cgtaacggc 47940  
 ttcagcgcg gctgggcaaa cgactctggc gactctggc cttcgaccac cgcaggttg 48000  
 agcgctgggt ggcgcatctc ctaccgacg tgctgaagc ggaggaccg agcgacacc 48060  
 ggcacatccg gtcgggtggc gcggatgac acatcgccat cgtcggtgcc gcctgccggt 48120  
 tccaggtgg ggatgaggg ctggagacat actggcgga tctggccgag ggcattggtg 48180  
 tcagcaccga ggtgccagc gaccggtggc gcgcggcgga ctggtacgac cccgatccg 48240  
 aggttccggg ccggaccat gtggccaagg tgccttccc ccgcatgtg cgcagctgg 48300  
 atcgccggtt cttcgccatt tcccctcgt aggcgatgag cctggaccg caacagcggc 48360  
 tgttgctgga ggtgagctgg gaggcgatc agcgcgctgg ccaggaccg atggcgctgc 48420  
 gcgagagcgc cacgggcgtg ttcgtgggca tgatcgggag cgagcacgcc gagcggtgc 48480  
 agggcctcga cgcagcgcg gcgttgtgt acggcaccac cggcaacctg ctcagcgctc 48540  
 ccgctggacg gctgtcgttc ttccgtgggtc tcgacggccc gacgatgacg gtggacacc 48600  
 cctgctcgtc gtcgctgggt gcgttgcaac tcgcctgcca gagcctgcga ttgggcgagt 48660  
 gcgaccaggc cctggccggc ggggtccagc tgcttttgtc gcccggtca ttcgtcgcg 48720  
 cgctcgccat gcgtttgctt tcgccagatg ggcggtgcaa gacgttctcg gccgctgcag 48780  
 acggcttttg gcgggcccag ggtggtgct tgggtggtg caagcggtc cgtgacgcgc 48840  
 agcgcgaccg cgacccatc ctggcggtgg tcaggagcac ggcgatcaac cacgatggcc 48900  
 cgagcagcgg gctcacggtg cccagcggtc ctgcccagca ggcgttgcta cgcagggcgc 48960  
 tggcgcaagc gggcggtggc ccggccgagg tcgatttctg ggagtgcac gggacgggga 49020  
 cagcgctggg tgaccgcata gaggtgcagg cgttggggc ggtgtacgg gggggccgc 49080  
 ccgcggagcg gccgtctgg ctgggcgtg tcaaggccaa cctcgccac ctggaggccg 49140  
 cggcgggctt ggccggcggt ctcaagggtg tcttggcgct ggagcacgag cagattcccg 49200  
 ctcaaccgga gctcgacgag ctcaaccgc acatcccgtg ggcagagctg ccagtggccg 49260  
 ttgtcccgag ggcggtcccc tggccgcgc gcgcgcgccc gcgtcgtgca ggcgtgagc 49320  
 ctttcggcct gagcgggacc aacgcgcatg tgggtgtgga ggaggcgcc gcggtggagc 49380  
 ctgtggcgcg ggcgccgag cgcgcagcg gcccggtgc ggcgatggag cagcgcgccg 49440  
 cggcgctgga tgcgcaggca gcccggctgc gggaccacct ggagaagcat gtcgagctt 49500  
 gcctcgccga tgtggcggtc agcctggcga gagggcgctc ttcggccgca caccggctgg 49560  
 cggtggccgc gagctcgcgt gagggcgctc cgtggggcg cctcgggcg cgcgagggc 49620  
 acacgcgcc cggccagggc tcgcagtggt tgggcatgg ccgaaagctc aaggtgggtc 49680  
 tcgtgtttcc cggccagggc ctggagggtt tgggcatgg ccgaaagctc atggccgaag 49740  
 agccggtctt ccggggcgcg ctggagggtt ggcgaccggc catcgaggcg gaagcgggc 49800  
 ggtcgctgct cggggagctc tccgcccagc aggccgcctc gcagctcggg cgcacgagc 49860



tggttcagcc ggtgctgttc gccatggaag tagcgctttc tgcgctgtgg cggctcgtggg 49920  
gagtggagcc ggaagcgggtg gtgggcccaca gcatgggcca ggttgcgggc ggcacagctgg 49980  
ccggcgcgct gtcgctcgag gacgcgggtg cgatcatctg ccggcgcgag cggctgctgc 50040  
ggcggtatcag cggtcagggg gagatggcgc tggctgagct gtcgctggag gaggccgagg 50100  
cggcgctgcg tggccatgag ggtcggctga gcgtggcggt gagcaacagc ccgcgctcga 50160  
ccgtgctcgc cggcgagccg gcggcgctct cggaggtgct ggccagccca tagcccgagc gtcgacccgc 50220  
gggtgttctg gcggcaggtg aaggtggacg tccgagccca gcgagcggct gcggtgccga 50340  
tgcgcgaaga gctgatcgcg gcgctgggag cgatccggcc gctcggtgcg agctactggg 50400  
tgcgctcgag ggtgacgggc ggggtgatcg cgggtccgga gcaagcgtg ctggagggtg 50460  
cggacaacct tcggcagccg gtgcgcttcg ctgcggcggc acccgatcct ggtgccgccc ctggacgaga 50520  
gccccgcgct gttcatcgag atgagccgc cggtgggctc gctgcggcga gggcaggacg 50580  
tccagacggc ggccgagcaa gggggcgctg cgctgtgggc gtccggctat ccggtgagct 50640  
agcgcgcgac gctgctggag gcgctgggga ggggtccgct gccgacctat ccctggcagc 50700  
gggctcggct gttccccgcg gcggcgagcg atgggtcgaa gccctcgctg cggcttcggc 50760  
acgagcggta ctggatcgag gacagcgtgc tgctcggggc tccattgctc gtctcggcgc 50820  
agcttcgcaa cggcgccacg gaccatccgc tgagcgacga gaggctatcc tacctttcgg 50880  
gacccggagc tcacttgtgg gccgtgttc gccagcggc cgctggtgct ggagcagctg gcgctcgagc 50940  
aacatagggt ccattggcga ggcagggcga gcctcgtgca agtgccctc agcgaagaag 51000  
ccgccggcgt agatctctat gaaggcggac gtcgtgagga ggcaggtagg agctgggtgc 51060  
gagccctcgc ggcctcattc caggtatcga gtcgtgagga ggtgggagcg ttgaaggaa 51120  
gtcccggtcg ggggacagtg ttagcgccc agagctcagc ggtgggagcg ctctatccgc 51180  
ggcacgccac ctcggtggga gattcaacgg gactatgggc gctcggagcg caggtgtggc 51240  
tgctcaacga gcacgccctc ggccgggtac gcttgccagg gggcgtggag agacatggca 51300  
tcggcacggg gagggtgctc gattcatccc gccttgttg atgcatgttt tcaggtgctg acagcgctgc 51360  
gcgcctaccg ggaattccat gagattcgga ggcgctgac ggcgctccac gaaccggatc 51420  
tcaccagccc tcccgcggtc cagggctccg gtgaatcaag cgggtgagtga cacctggctg tgggacgccg 51480  
cgctggacgg tggacggcgc cagagcgca gcgtgcccg cgacctgggtg ctccggcagct 51540  
tccatgcgaa gtgggaggtc atggagcgcc tcgcgcaggc gtacatcatc ggcactctcc 51600  
gcatatggaa cgtcttctgc ctctgtcgtc tacaggaagg gacatcacac agctcagag 51660  
ggcttcaaat ctctgtcgtc gacggagagc attttgtgag atggatggaa caccttgtcg 51720  
cgatcggcat ccttgtaggg gaggaggccg ggagggtgtt cgcggagctc gatagacgag 51780  
ctgatttgge caagtttgcc ggtggctcgt tgcgcggtgt tggagatcgg accggacctc 51840  
ttgagtgggt cgctcgagat cctcttccct aacggcatcg ggcggagctc agcagggagc 51900  
cgcccatcgc ccgttactcg gtcgggaatg ttcagcatct ttgctgcctg accggacgga 51960  
tggttagcacc gtcgggaatg gtcgagcctg gtcgagatcg ggcggcgagg ggcgcgacca 52020  
ccgcgcgctg ccttgcctgc gcgagcagct gagccagata gccagggata cgcacatcag 52080  
ctccgctctt ccttgcctgc gcgagcagct gcccgatat tgggtgctggg cgagggcaca 52140  
gcattctgga tgtcgaccag ggccaatgtc gctcgcgccc ggaggccttc aggggtggca 52200  
tcacgtcgc ggcgaatgtc gctcgcgccc ggattgattg aggggtggca gaagtagcaa gatgatctc 52260  
tctgtcgtt tatcaccacg cctgctcgtg cctggtgta gatctccggc ggggatcctc 52320  
gtatcgacca tccgctcgt cctgctcgtg ccaggcgacg ggcagtagcag gagccgctt 52380  
ttgaggacgc cgtgagctg gcgcgcgccc ggcagtagcag ggcagtagcag ggcagtagc 52440  
tgatcctctc ggcgcgccc gatggcggtg accggtgagc acttccaccg ggcagtagc 52500  
cgaccgaatc ggcgcgccc gatggcggtg accggtgagc accggtgagc accggtgag 52560  
tcgtccatcg gatggcggtg accggtgagc accggtgagc accggtgag 52620  
gggtccacgg tccgctcgt cgaagacacc gggcaggtcg ggcagtagcag ggcagtagc 52680  
tgctcctgtt agctcgagge ttctgcttcc ataccagag gaccagggc gaccagggc 52740  
aatggcagcg gggcttggtc cgtgctgag gaccagggc gaccagggc gaccagggc 52800  
aaggcgaggg cgaggcgtgc gtgcgcgtca tcgcgggtac ggcagtagcag ggcagtagc 52860  
cggggctgta tcaagtcat ccggcgcgag tagtgcatat gtcgagtagc ggcagtagc 52920  
cattcggcga gaggcgagc ggcgagtagc ttcagggcga ggcagtagc ggcagtagc 52980  
cagcagggga gaggcgagc ggcgagtagc ttcagggcga ggcagtagc ggcagtagc 53040  
gcgcgcttcc tctggtgcag catgcggtgg ggcagtagc ggcagtagc ggcagtagc 53100  
ggctcttgac ccgcgcccgt gtggggcctc cgtgaaccg gcgagtagc ggcagtagc 53160  
aggcgccggt cgtgaaccg gcgagtagc ggcagtagc ggcagtagc ggcagtagc 53220  
cgctcgtgga cgtgaaccg gcgagtagc ggcagtagc ggcagtagc ggcagtagc 53280  
tcggggcgag cgacagagag gaccaggtcg ggcagtagc ggcagtagc ggcagtagc 53340  
gctcgtgceg gagctcctt ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53400  
gcagctatgt gatcaccgat ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53460  
tgatgcaggg ggcccgcct ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53520  
ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53580  
ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53640  
ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53700  
ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53760  
ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc ggcagtagc 53820



gggatgccct cccgtccatg gccgaggctg gcgaggaggt gcagatcgtg gaggccgacg 53880  
 tggctcggcg cgacgatgtc gctcggctcc tctcgaagat cgaaccgtcg atgccgcgcg 53940  
 ttcgggggat cgtgtacgtg gacgggacct tccagggcga ctccctcgatg ctggagctgg 54000  
 atgcccgctg cttcaaggag tggatgtatc ccaaggtgct cggagcgtgg aacctgcacg 54060  
 cgctgaccag ggatagatcg ctggacttct tctcctgtgta ttcctcgggc acctcgcttc 54120  
 tgggcttgcc aggacagggg agccgcgcgc ccggtgacgc cttcttggac gccatcgcg 54180  
 atcaccgggt caaggtgggc cttacagcga tgagcatcaa ctggggattg ctctccgaag 54240  
 catcatcgcc ggcgaccccg aacgacggcg gcgacggct cgaataccgg gggatggaag 54300  
 gcctcacgct ggagcagggg gcggcggcgc tccggcgctt gctcgacga cccagggcg 54360  
 aggtaggggt gatcgggctg aatctgcgc agtggttggg gttctatccc aacggggccc 54420  
 gattggcgct gtggcgggag ctgctgaagg agcgtgacgc cgccgaccga ggcgcgcg 54480  
 acgctcgaa cctgcgcgag gcgctgcaga gcgccaggcc cgaagatcgt cagttgattc 54540  
 tggagaagca cttgagcgag ctggtggggc ggggctgctg ccttccgcgc gagaggatcg 54600  
 agcggcacgt gccgttcagc aatctcggca tggactcgct gataggcctg gagctccgca 54660  
 accgcatcga ggccgcgctc ggcatcaccg tgccggcgac cctgctatgg acctacccta 54720  
 acgtagcagc tctgagcggg agcttgctag acattctgtt tccgaatgcc ggccgacccc 54780  
 acgctccggc caccgagcgg gagaagagct tcgagaagca tgccgcagat ctcgaggctc 54840  
 tgcggggcat gacggacgag cagaaggacg cgttgctcgc cgaaaagctg gcgcagctcg 54900  
 cgcagatcgt tggtagtaaa gggaccgagg gagtatggcg accacgaatg ccgggaagct 54960  
 tgagcatgcc cttctgctca tggacaagct tgcgaaaaag aacgcgtctt tggagcaaga 55020  
 gcggaccgag ccgatcgcca tggctgcgcg tcccccggcg gagcggacac 55080  
 tccggaggca tcttgggagc tgctcgactc agggccgagc gcggtccagc cgctcgaccg 55140  
 gcgctggggc ctggctggcg tccatcccag cgaggaggtg ccgcgctggg ccggactgct 55200  
 caccgaggcg gtggacggct tcgacgccgc gttctttggc acctcgctc gggaggcgcg 55260  
 gtcgctcgat cctcagcaac gcctgctgct ggaggtcacc tgggaagggc tcgaggacgc 55320  
 cggcatcgca ccccagtcct tcgacggcag ccgcaccggg gtgttcttgg gcgcagcg 55380  
 cagcgactac tcgcataccg ttgcgcaaca gcggcgcgag gagcaggacg catacgacat 55440  
 caccggcaat acgctcagcg tcgccgcgcg acggttgtct tatacgctag ggctgcaggg 55500  
 accctgctg accgtcgaca cggcctgctc gtcgctcgctc gtggccatcc acctgctg 55560  
 ccgcagctg cgcgctcgcg agagcgatct agcgtggcg ctgctgcccg acatgctcct 55620  
 ttcgtccaag acgatgataa tgctggggcg catccaggcg ctgtcgcccg atggccactg 55680  
 ccggacattc gacgcctcgg ccaacgggtt cgtccgtggg gagggtgctg gtatggtcgt 55740  
 gctcaaaccg ctctccgacg cccagcgaca cggcgatcgg atctgggctc tgatccggg 55800  
 ttcggccatg aatcaggatg gccggtcgag aggttgtatg gcaccaatg tgctcgctca 55860  
 ggagcgctc ttgcgcgagg cgctgcagag cgctcgctc gacgccgggg ccacggtta 55920  
 tgcgagacc cacggaacgg ggacctcgct cggcgaccgc atcgaggtcg aggcgctcg 55980  
 tgccgtgttg gggccggcgc gggccgatgg gagccgctgc gtgctggggc cagtgaagac 56040  
 aaacctcggc cacctggagg gcgctgcagg cgtggcggtt ttgatcaagg cggcgctggc 56100  
 tctgcaccac gaactgatcc cgctggcgac cctcagcggc ccgtggccgc cgcggatccg 56160  
 gatcgagggg acccgctcg gcgggggtga gcgcgttcgg cctcagcggc accaactcc 56220  
 accgcgcttc cggggccacg tgcctgcacc ggcgacgcgc gggcgctcag cggagctttt 56340  
 ggaggaggcg gcgaagagcg ccgcgcgcgt ggacgcacag ggcgcgcggc tctcagcgca 56400  
 ggtgctgtcg taccggagc agaggtctcg ttcagcctgg tatcgacgcg 56460  
 catcgccgct tagccgatg gagcaccggc tcgcggtggc ggcgacctcg cgcgagggcg 56520  
 gctggagggt gcggcgaggg ggcagacccc ggcaggcgcg gcgcgcggca gggccgctt 56580  
 ctgcggcggc aagctcgctt cctgttgcgc cgggcagggc gcgcaggtgc cgggcggtgg 56640  
 ccgtgggttg tgggaggcgt ggccggcggt cgcgagacc ttcgaccggt gcgtcacgct 56700  
 cttcgaccgg gagctccatc agcgtctctg cgaggtgatg tgggcccagc cgggcagcag 56760  
 caggctcgct tgctggacc agacggcggt caccagccg gcgctctttg cgtggagta 56820  
 cgcgctggcc gcgctcttcc ggtcgtgggg cgtggagccg gagctcgctg ctggccatag 56880  
 cctcggcgag ctgggtggcg cctgcgtggc ggtgtgttcc tccctcgagg acgcccgtg 56940  
 cttggtggtc gcgcgcggcc ggttgatgca ggcgctgcgc gccggcgggc cgttgggtgc 57000  
 gatcgccgcg cgggagggcg acgtggctgc ccggtggcg ccgcacgcag cgttgggtgc 57060  
 gatcgccgca gtcaatgggc cggagcaggt ggtgatcgcg ggcgcggaga aattcggtga 57120  
 gcagatcgcg gcggcgctcg cggcgcgggg ggcgcgaacc aaaccgctgc atgtctcgca 57180  
 cgcgttccac tcgccgctca tggatccgat gctggaggcg ttcggcgggg tgactgtgac 57240  
 ggtgacgtac cggcggcctt cgatcgcgct gttactgggt gcgtcacgcg cgagaggcgg 57300  
 cgatgaggtg agcgcgcggg gttactgggt gcgtcacgcg cgcgagggcg 57360  
 ggacggagtg aaggcgctgc acgcggccgg tcggggcctc ttcgctcgagg tggggccgaa 57420  
 gccgacgctg ctccggcctt tgccggcctg cctgcccgat gccaggccgg tgctgctccc 57480  
 agcgtcgcg gcggggcggt acgaggctgc gaggcgctga gaggcgctgg gtgggttctg 57540  
 ggtcgtcggt ggatcggtca cctggtcggg tgcttccct tcgggcggac ggcgggtacc 57600  
 gctgccaacc tatccctggc agcgcgagcg ttactggatc gaagcgccgg tcgatcgta 57660  
 ggcggacggc accggccgtg ctccggcggg gggccacccc cttctgggtg aagtcttttc 57720  
 cgtgtcgacc catgccggtc tgcgcctgtg ggagacgacg ctggaccgaa agcggctgcc 57780

gtggctcggc gagcaccggg cgcaggggga ggtcgtgttt cctggcgccg ggtacctgga 57840  
 gatggcgctg tcgtcggggg ccgagatctt gggcgatgga ccgatccagg tcacggatgt 57900  
 ggtgctcatc gagacgctga ccttcgcggg cgatacggcg gtaccgggcc aggtgggtgac 57960  
 gaccgaggag cgaccgggac ggctgcgggt ccaggtagcg agtcgggagc cgggggaacg 58020  
 tcgcgcgccc ttccggatcc acgcccgcgg cgtgctgcgc cggatcgggc gcgtcgagac 58080  
 cccggcgagg tcgaacctcg ccgcccctgcg cggccggctt catgccgcgc tgcccgcgtgc 58140  
 ggctatctat ggtgcgctcg ccgagatggg gcttcaatac ggcccggcgt tcgggggggct 58200  
 cgcccgagctg tggcgggggt agggcgaggc gctgggcagg gtgagactgc ctgaggccgc 58260  
 cggctccgcg acagcctacc agctgcatcc ggtgctgctg gacgcgtgcg tccaaatgat 58320  
 tgttgccgcg ttcccgatc gcgatgaggc gacgccgtgg gcgccgggtg aggtgggctc 58380  
 ggtgcggtcg ttccagcggg ctccctgggga gctatggtgc catgcgcgcg tcgtgagcga 58440  
 tggccaacag gcctccagcc ggtggagcgc cgactttgag ttgatggacg gtacggggcg 58500  
 ggtggtcgcc gagatctccc ggctggtggt ggagcggctt tgaggagccc gggcgctcg gtgggcccga 58560  
 cgacgcagac gactggttcc tggaggtgga ttgggagccc gggctcgggc gctcgttggtg 58620  
 gatcacagcc ggccgggtggc tgctgctcgg cgagggtggt gggctcgggc gctcgttggtg 58680  
 ctccggcgtg aaggccgcgc gccatgtcgt cgtccacgcc gcgggggacg acacgagcac 58740  
 tgcaggaatg cgcgcgctcc tggccaacgc gttcgacggc caggccccga cggccgtggt 58800  
 gcacctcagc agcctcgacg gggcgggcca gctcggcccg gggctcgggg cgaggggcg 58860  
 gctcgacgcg ccccgaggcc cagatgtcga tgccgatgcc ctcgaaatcg cgctgatgcg 58920  
 tggttgcgac agcgtgctct ccctggtgca agcgtggtc ggcagggacc tccgaaacgc 58980  
 gccgcggctg tggctcttga cccgcggggc tcaggcgccc gcccgccggc atgtctccgt 59040  
 ggtgcaagcg ccgctgttgg ggctggggcg caccatgcc ctgagcaacg ccgagctcg 59100  
 ctgtatcagc gtcgacctgc atccagccga gcctgaagg gaagccgatg cttfgctggc 59160  
 cgagctactt gcagatgat gcgagggag ggtcgcgctg cgcggtggcg accggtcgt 59220  
 tgccggtcgc gtccaccggc tgcccagcgc tcagcgccgg gagaaggctg agcccgcgg 59280  
 tgacaggccg ttccggctag agatcgatga accggcgcg ctggaccaac tggtgctccg 59340  
 agccacgggg cggcgctcgc atggtccgg agctggcgtt gggcgctgct cccaatgatc 59400  
 gctcgactcc atcgacatcc ccgttggtgc tcggaagcga gtgcgcggg cgcacgtcg 59460  
 agaaatcgag ccgttggtgc tgggcccagc ggtgatcgcc ctgcccggcg gagtatttg 59520  
 gggcggtgaac ggccttggtg ccacgtggt gttgctcgg cctctggggc tctcgccgac 59580  
 taccatgtc accacgtcgg ccagctggt gacggcctgg gacggcctgg tacgccctcg 59640  
 cgaggcgccc gcgatgcccc ggcgtatgt ccagtcggg gcccgtggtg tccgtctttg 59700  
 ccacctgcag gcgggggagc ggggcccagc ggcgtgacg cctctggggc acaaggtcgc 59760  
 cgcggtgcga tgggcccagc gcgtgggccc cggttacgtg agcgattccc gctcgggccc 59820  
 gaaccgtgcc ttacctggag catgggacg catgggacgt ggggacgtcg tgctcgactc 59880  
 gttcgtcaca gacgtgcag acaagagcct cactggcctg cgcgcctgtg gtcgccttgt 59940  
 gctttcgggg gagcgcacg actgcgccga cagcgagcct gggctgccgc cgctcctacg 60000  
 gaagctgggg aggcgcgac tggacttgcg ggggaatgat ggtgccatca gctcggcagc 60060  
 gaatttttcc ttctcgcagg tggacttgcg ggtgccatca ggtgccatca gctcggcagc 60120  
 ccgtgctcgc ctcgacgagc tggtcggggt gatccctcac gccaccgcgc gtcgagacct 60180  
 tgcggggttg cgcgttgggc ggggggacg gcaaggacag catctcggga agctcgtgct 60240  
 tcgcccagcc gaggcattcc ggaggatggc cgtcgggccc gaatccagcg tcgcccgtcg 60300  
 cacgctggac gacccggagg tgcggatccg cctgggtggc ctcggtctgc gctgggccc 60360  
 cgcggaagcc acctacctg tgaccggcgg tctgggtggc gggcgctcgc gtgcccagc 60420  
 atggctggcc gagcggaagt cgggggacg ggtgctggtg ggcgcgcgcg tcacgggtgg 60480  
 cgcagagcag cgagccgcg tggcggcgct ggaggcccac ggcgcgcgcg ttaccgcgtc 60540  
 gaaagcggac gtcgccgat ggtcacagat ggcagggtct gtggatgacg ggctgctgat 60600  
 ggggatgccg ctgcccgggt tcgtgcatgc tccgcacggt aaggtccagg gggccttgca 60660  
 gcagcagact ccggcgcggt tccgcacggt aagcgctct tctctcttc gtgctgtacg 60720  
 cttgcacacg ctgacacggt aagcgctct tctctcttc gtgctgtacg cccaacgcgt 60780  
 tgggcttttc ggctcgccag gccagggcaa ctatgccgca gccagcgct agcatcgact 60840  
 cctttcgcac caccgaaggg cgcagggccc gcccggcgct agcatcgact ggcgggcaga 60900  
 cacggagggt gggatggccg ttgcgcaaga aaacgggtgg gcgcggcaga tctctcgcgg 60960  
 gatcgggggc atcaccctcg atgagggtct tcgcggcag ggcgcgttg tcgagggtga 61020  
 tcgctgcag acgggggtga taccgatcac tccgcggcag tgggtggagt tctacccggc 61080  
 aacagcggcc tcacggagg gtgcgcggct ggtgaccacg cagcgcgcg gctgagccga 61140  
 gaccgcccgg gatcgggacc tgctcgaaca gcttgctcg gctgagccga gcgcgcggc 61200  
 ggggctgctg caggacgtcg gtcgcgtgca ggtctcgcac gtgctgcgtc tccctgaaga 61260  
 caagatcgag gtggatgccc cgctctcgag catgggcatg gactcgctga tgagcctgga 61320  
 gctgcgcaac cgcacagagg ctgcgctggg cgtcgccgcg cctgcagcct tgggggtggac 61380  
 gtaccaacag gtagcagcga taacgcgctg gctgctcgac gacgccctcg tcgtccggct 61440  
 tggcgccggg ctcggacagc acgaatcgac ggcgagcgcc ggttcggtcg tccacgtcct 61500  
 ccgcttcgt cctgctgca agccgcgggc tegtctctt tgttttcacg gttctggcgg 61560  
 ctgcccag ggcttccgtt cctggtcgga gaagtctgag tggagcgatc tggaaatcgt 61620  
 ggcatgtgg cacgatcgca gcctcgctc cgaggacgcg cctggtaaga agtacgtcca 61680  
 agaggcgccc tcgctgattc agcactatgc agacgcaccg tttgcgttag tagggttcag 61740

cctgggtgtc	cggttcgtca	tggggacagc	cgtggagctc	gccagtcggt	ccggcgcacc	61800
ggctccgctg	gccgtcttca	cggttgggagg	cagcttgatc	tcttcttcag	agatcacccc	61860
ggagatggag	accgatataa	tagccaagct	cttcttccga	aatgccgcgg	gtttcgtgcg	61920
atccacccaa	caagtccagg	ccgatgctcg	cgcagacaag	gtcatcacag	acaccatggg	61980
ggctccggcc	cccggggact	cgaaggagcc	gcccgtgaag	atcgcgggtcc	ctatcgtcgc	62040
catcgccggc	tccgacgatg	tgatcgtgcc	tccgagcgac	gttcaggatc	tacaatctcg	62100
caccacggag	cgcttctata	tgcattctct	tcccggagat	cacgaatttc	tcgtcgatcg	62160
agggcgcgag	atcatgcaca	tcgtcgactc	gcattctcaat	ccgctgctcg	ccgcgaggac	62220
gacgtcgta	ggccccgcgt	tcgaggcaaa	attgtaggag	cctccctcgg	gcgcgcgaga	62280
tggttgggag	cagcgtgggc	gctggcggcc	ggcggcaggc	cgccgaggcg	catgagcctt	62340
cctggacggt	tgcagtatag	gagattttat	gacacaggag	caagcgaatc	agagtggagac	62400
gaagcctgct	ttcgacttca	agccgttcgc	gcctgggtac	gcccaggacc	cgttccccgc	62460
gatcgagcgc	ctgagagagg	caacccccat	cttctactgg	gatgaaggcc	gctcctgggt	62520
cctcacccga	taccacgacg	tgctggcggt	gttccgcgac	gaacgcttcg	cggtcagtcg	62580
agaagagtgg	gaatcgagcg	cggagtactc	gtcggccatt	cccagactca	gcgatatgaa	62640
gaagtacgga	ttgttcgggg	tgccgcggga	ggatcacgct	cgggtccgca	agctcgtcaa	62700
cccgtcggtt	acgtcacgcg	ccatcgacct	gctgcgcgcc	gaaatacagc	gcaccgtcga	62760
ccagctgctc	gatgctcgct	ccggacaaga	ggagtccgac	gttgtgcggg	attacgcgga	62820
gggaatcccc	atgcgcgcga	tcagcgctct	gttgaagggt	ccggccgagt	gtgacgagaa	62880
gttccgtcgc	ttcggctcgg	cgactgcgcg	cgcgctcggc	gtgggtttgg	tgccccaggt	62940
cgatgaggag	accaagaccc	tggtcgcgtc	cgtaaccgag	gggctcgcgc	tgctccatga	63000
cgctcctcat	gagcggcgca	ggaacccgct	cgaaaatgac	gtcttgacga	tgctgcttca	63060
ggccgaggcc	gacggcgaca	ggctgagcac	gaaggagctg	gtcgcgctcg	tggtgtcgat	63120
tatcgctgct	ggcaccgata	ccacgatcta	ccttatcgcg	ttcgctgtgc	tcaacctgct	63180
gcggtcgcgc	gaggcgctcg	agctggtgaa	ggccgagccc	gggtcatga	ggaacgcgct	63240
cgatgaggtg	ctccgcttcg	acaatatctc	cagaatagga	actgtgcgtt	tcgccaggca	63300
ggacctggag	tactgcgggg	catcgatcaa	gaaaggggag	atggtctttc	tcctgatccc	63360
gagcgcctcg	agagatggga	ctgtattctc	caggccagac	gtgtttgatg	tgcgacggga	63420
cacgggcgcg	agcctcgcgt	acggtagagg	cccccatgtc	tgccccgggg	tgctcccttc	63480
tcgcctcgag	gcggagatcg	ccgtgggcac	catcttccgt	aggttccccg	agatgaagct	63540
gaaagaaact	cccgtgtttg	gataccaccc	cgcgctccgg	aacatcgaat	cactcaacgt	63600
catcttgaag	ccctccaaag	ctggatagct	cgcgggggta	tcgcttcccc	aacctcatte	63660
cctcatgata	cagctcgcgc	gcgggtgctg	tctgccgcgg	gtgcgattcg	atccagcgga	63720
caagcccatt	gtcagcgcgc	gaagatcgaa	tccacggccc	ggagaagagc	ccgtccgggt	63780
gacgtcggaa	gaagtgcggg	gcgcgcgcct	gggagcgcga	agctcgcctc	ttcgcgctca	63840
gcacgcgcgt	cgatcatgtcc	ggccctgcac	ccgcgcggag	gagccgcccc	ccctgatgca	63900
cgccctcacc	gagcggcagg	ttctgctctc	gctcgtcgcc	ctcgcgctcg	tcctcctgac	63960
cgcgcgcgcc	ttcggcgagc	tcgcgcggcg	gctgcgccag	cccaggtgct	tcggcgagct	64020
cttcgggggc	gtggtgctgg	gcccgtccgt	cgctggcgcg	ctcgtcctcg	ggttccatcg	64080
agtcctcttc	caggatccgg	cggtcggggg	cggtctctcc	ggcatctcct	ggataggcgc	64140
gctcgtcctg	ctgctcatgg	cggtatcgca	ggtcgatgtg	agcatcctgc	gcaaggaggc	64200
gcgccccggg	gcgctctcgg	cgctcggcgc	gatcgcgcgc	ccgctgcgca	cgccggggcc	64260
gctggtgcag	cgcagtcagg	gcgcgttcac	gtgggatctc	gacgtctcgc	cgcgacgctc	64320
tgcgcaagcc	tgagcctcgg	cgccctcgcc	tacacctcgc	cggtgctcgc	tccgcccgcg	64380
gacatccggc	cgcccgcgcg	ggcccagctc	gagccggact	cgccggatga	cgaggccgac	64440
gaggccgacg	aggcgctccg	cccgttccgc	gacgcgatcg	ccgcgtactc	ggaggccgtt	64500
cggtggggcg	aggcggcgca	gcggccgcgg	ctggagagcc	tcgtgcggct	cgcgatcgtg	64560
cggctgggca	aggcgctcga	caaggtccct	tcgcgcaca	cgacggccgg	cgctccccag	64620
atcgccggca	gactccagaa	cgatgcggtc	tggttcgatg	tcgccgcccg	gtacgcgagc	64680
ttccgcgcgg	cgacggagca	cgcgctccgc	gacgcggcgt	cgcccatgga	ggcgctcgcg	64740
gccggcccgt	accgcggatc	gagccgcgtg	tccgctgccc	taggggaggt	tcggggggag	64800
gcggcgcgcc	ttcaccccg	ggaccgtgta	cccgcgtccg	accagcagat	cctgaccgcg	64860
ctgcgcgcag	ccgagcgggg	gctcatcgcg	ctctacactg	cgttcgcccc	tgaggagtga	64920
gcctctctcg	ggcgacggcg	agcggcgggc	tgccggtggt	tcctctcttc	caaccatgac	64980
cggagccgcg	ctcgggtccgc	gcagcggcta	gcgcgcgtcg	cggcagagat	cgctggagcg	65040
acaggcgacg	accgcggcga	gggtgtcgaa	cggattgccc	cagccctcat	tgccgatccc	65100
ctccagacac	tcgttcagct	gcttggcgct	gatgcgcgct	gggcactcgc	cgaaggctcag	65160
ctcgtcgcgc	cactcggatc	ggatcttgtt	cgagcacgcg	tccttgctcg	aatactcccc	65220
gtcttgtccg	atgttgttgc	accgcgcctc	gcggtcgcac	cgcgccgcca	cgatgctatc	65280
gacggcgctg	ccgactggca	ccggcgccct	gccctgcgcg	ccaccggggg	tttgcgcctc	65340
cccgcctgac	cgcttttcgc	cgccgcacgc	cgcgagcagg	ctcattcccc	acaccgagat	65400
caggccccag	accagcttcc	cagcaatctt	ttgcatggct	tcccctccct	cacgacacgt	65460
cacatcacag	actctccgct	cgctcgtcg	gttcgacagc	cggcgacggc	cacgagcaga	65520
accgtccccg	accagaacag	ccgcatgcgg	gtttctcgca	acatgccccg	acatccttgc	65580
gactagcgtg	cctccgctcg	tgccgagatc	ggctgtcctg	tgcgacggca	atatcctgcg	65640
atcgccgggg	caggaggtac	cgacacgggc	gccggggcgg	aggtgccgcc	acgggctcga	65700

```

aatgtgctgc ggcaggcgcc tccatgcccc cagccgggaa cgcggcgccc ggccagcctc 65760
gggggtgacgc cgcaaacggg agatgctccc ggagaggcgc cgggcacagc cgagcgccgt 65820
caccaccgtg cgactcgtg agtccagct cctcggcata gaagagaccg tcaactcccg 65880
tccgtgtagg cgatcgtgct gatcagcgcg ttctccgctt gacgcgagtc gagccgggta 65940
tgctgcacga caatgggaac gtccgattcg atcacgctgg catagtccgt atcgcgcggg 66000
atcggtcggg gttcgggtcag atcggtgaac cggacgtgcc ggggtgcgct cgctgggacg 66060
gtcaccgggt acggccccgc ggggtcggcg tcgctgaagt agacgggtgat ggcgacctgc 66120
gcgtcccggt ccgacgcatt cccggcctgg gctcgtcat ggctcgtcat ctgcggctcg 66180
ggtcggttgc tccggcctgg gatgtagccc tctgcgattg cccagcgctg ccgcccgatc 66240
ggcttctcca tatgtctccc ctgctggctc ctctttggct gcctccctct gctgtccagg 66300
agcgacggcc tcttctcccc acgctcggg ggatccatgg ctgaggatcc tcgcccagcg 66360
ctccttgccg accggcgcgc cgagcgccga cgggctttga cgggctcgtg acgacgcga ccggacacgt 66420
gatgcccggg cgacaggccc gcccgcgtc tgatcccgat cgtgacatcg cgacgtccgc 66480
cggcgccctc gcaggccggc ctgagcggtg cgcggtcatg gtcgtcctcg cgtcaccgcc 66540
acccgcccgt tccatcccca ccgcggaacg acgcttgctc aaaccgcggc gagacggcgg 66600
ggcggctgtg gtaccggcca gcccggagcg gaggccccag agggacagtg ggtccgccgt 66660
gaagcagtga ggcatcgag gtggcagatg aaacacgttg acacgggccg acgagtcggc 66720
cgccggatag ggctcacgct cggctcctc cggagcatgg cgctcgccgg ctgtggcgcc 66780
ccgagcgaga aaatcgtgca gggcacgcgg ctgcgccccg gcgcccgatg gcacgtcgcc 66840
gcccagctcg accccgacgc cgcgaccacg cgggtggcgg tggacgtcgt tcacctctcg 66900
ccgcccagag gcacagggc cggcagcgag cgggttcgtc tctggcagcg tccgagctcc 66960
gagtcctcgt ggcaacgggt cggagtgctc gactacaacg ctgccagccg aagaggcaag 67020
ctggccgaga cgaccgtgcc cgatgccaac ttccagctgc tcatcaccgt cgagaagcag 67080
agcagccctc agtctccatc ttctgccgcc gtcacggggc cgacgtccgt cgggtaacat 67140
cgcgctatca gcagcgctga gcccgccagc agggccccaga gccctgcctc gatcgccctc 67200
tccatcatat catccctgcg tactcctcca gcgacggccg cgtcgaagca accgcccgc 67260
cggcgccggt ctacgtgcgc gacaggagag cgtcctggcg cggcgcccca tccgagccaa 67320
ggatcggcgg agcatggaga aagaatcgag gatcgcgatc accggcagct cggcgatgct 67380
cgtggcgatc gcggcggtca agttcatcgc cgccgccgtg gctcctcctc tgctcggcaa 67440
ctccgagggc gtgcaactccc tcgtcgatac tgacagcggg ctcctcctcc tgctcggcaa 67500
gcaccggagc gcacgcccgc ccgacgccga gcatccgttc ggccacggca aggagctcta 67560
tttctggacg ctgatcgtcg ccatcatgat cctcgcggcg ggcggcgccg tctcgatcta 67620
cgaaggatc ttgcacctct tgcacccgcg ccagatcgag gatccgacgt ggaactacgt 67680
cgtcctcggc gcagcgcccg tcttcgaggg gacgtcgctc atcatctcga tccacgagtt 67740
caagaagaag gacggacagg gctacctcgc ggcgatgcgg tccagcaagg acccgacgac 67800
gttcacgatc gtccctggag actccgcccg gctcgcgggg ctcaccatcg ccttctcctg 67860
cgtctggctc gggcacccgc tcgcccggcg cctcctcgac ggcgcccgtg cgatcggcac 67920
cgccctcggt ctcgcccggc agctcctcgc cgcgatccgc gcgctcgcca ggcgagatcc 67980
gggggagagc gcggaacggg ggcgggtcgt cgtgcacttc ggtccgcacg aagtcctggt 68040
tggcgtgtcg gcggtggggc acgcccgcgt caccgctcc cggggtcggg aggcgatcga 68100
cgtgctgcgc atcgagttcg acgcccgcgt ggagcgagcg acccgacgtg aagcacatct acgtcgaggc 68220
gcgcatcgag acccgatac cgagggcggt acgcccgtg gagagaccgc ggcggccctc 68280
caggctcgct cgcggcgccc gggctcaggt gggcctcgca gcagggcgcg cctggcgggc 68340
aaaccgtgca gacgtcgctc ttgcagcgga ggtacgttgg ttgcaagtcg tcacgcccga 68400
tcgagaggtc cggcagcgcc gggagcccgg gtcagagga agggcagggc gcacgaaggc 68460
caggcttcga gggggcgac gtcagagga agggcagggc gcatggggcg atgctcgggc 68520
ggcgagatga cggctggcgt cgcggcctcc cgcggcgccg cgcgcttcgc gcccgctccc 68580
agcgcggtcg ctgcgcgat ctgcggcgcc gcccgtcat cgcctcctg tccctcgccg 68640
gcggcgccag catggcggtc gtctcgctgt tccagctcgg gatcatcgag cgcccgccc 68700
atcctccgct tccagggttc gattcggcca aggtgacgag ctccgatata 68750

```

&lt;210&gt; 2

&lt;211&gt; 1421

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 2

Val Ala Asp Arg Pro Ile Glu Arg Ala Ala Glu Asp Pro Ile Ala Ile  
 1 5 10 15

Val Gly Ala Ser Cys Arg Leu Pro Gly Gly Val Ile Asp Leu Ser Gly  
 20 25 30

Phe Trp Thr Leu Leu Glu Gly Ser Arg Asp Thr Val Gly Arg Val Pro  
 35 40 45

Ala Glu Arg Trp Asp Ala Ala Ala Trp Phe Asp Pro Asp Pro Asp Ala  
 50 55 60  
 Pro Gly Lys Thr Pro Val Thr Arg Ala Ser Phe Leu Ser Asp Val Ala  
 65 70 75 80  
 Cys Phe Asp Ala Ser Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Arg  
 85 90 95  
 Met Asp Pro Ala His Arg Leu Leu Leu Glu Val Cys Trp Glu Ala Leu  
 100 105 110  
 Glu Asn Ala Ala Ile Ala Pro Ser Ala Leu Val Gly Thr Glu Thr Gly  
 115 120 125  
 Val Phe Ile Gly Ile Gly Pro Ser Glu Tyr Glu Ala Ala Leu Pro Gln  
 130 135 140  
 Ala Thr Ala Ser Ala Glu Ile Asp Ala His Gly Gly Leu Gly Thr Met  
 145 150 155 160  
 Pro Ser Val Gly Ala Gly Arg Ile Ser Tyr Ala Leu Gly Leu Arg Gly  
 165 170 175  
 Pro Cys Val Ala Val Asp Thr Ala Tyr Ser Ser Ser Leu Val Ala Val  
 180 185 190  
 His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Ser Thr Ala Leu  
 195 200 205  
 Ala Gly Gly Val Ser Leu Met Leu Ser Pro Ser Thr Leu Val Trp Leu  
 210 215 220  
 Ser Lys Thr Arg Ala Leu Ala Arg Asp Gly Arg Cys Lys Ala Phe Ser  
 225 230 235 240  
 Ala Glu Ala Asp Gly Phe Gly Arg Gly Glu Gly Cys Ala Val Val Val  
 245 250 255  
 Leu Lys Arg Leu Ser Gly Ala Arg Ala Asp Gly Asp Arg Ile Leu Ala  
 260 265 270  
 Val Ile Arg Gly Ser Ala Ile Asn His Asp Gly Ala Ser Ser Gly Leu  
 275 280 285  
 Thr Val Pro Asn Gly Ser Ser Gln Glu Ile Val Leu Lys Arg Ala Leu  
 290 295 300  
 Ala Asp Ala Gly Cys Ala Ala Ser Ser Val Gly Tyr Val Glu Ala His  
 305 310 315 320  
 Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ile Gln Ala Leu Asn  
 325 330 335  
 Ala Val Tyr Gly Leu Gly Arg Asp Val Ala Thr Pro Leu Leu Ile Gly  
 340 345 350  
 Ser Val Lys Thr Asn Leu Gly His Pro Glu Tyr Ala Ser Gly Ile Thr  
 355 360 365  
 Gly Leu Leu Lys Val Val Leu Ser Leu Gln His Gly Gln Ile Pro Ala  
 370 375 380  
 His Leu His Ala Gln Ala Leu Asn Pro Arg Ile Ser Trp Gly Asp Leu  
 385 390 395 400

Arg Leu Thr Val Thr Arg Ala Arg Thr Pro Trp Pro Asp Trp Asn Thr  
 405 410 415  
 Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Met Ser Gly Thr Asn Ala  
 420 425 430  
 His Val Val Leu Glu Glu Ala Pro Ala Ala Thr Cys Thr Pro Pro Ala  
 435 440 445  
 Pro Glu Arg Pro Ala Glu Leu Leu Val Leu Ser Ala Arg Thr Ala Ser  
 450 455 460  
 Ala Leu Asp Ala Gln Ala Ala Arg Leu Arg Asp His Leu Glu Thr Tyr  
 465 470 475 480  
 Pro Ser Gln Cys Leu Gly Asp Val Ala Phe Ser Leu Ala Thr Thr Arg  
 485 490 495  
 Ser Ala Met Glu His Arg Leu Ala Val Ala Ala Thr Ser Arg Glu Gly  
 500 505 510  
 Leu Arg Ala Ala Leu Asp Ala Ala Ala Gln Gly Gln Thr Ser Pro Gly  
 515 520 525  
 Ala Val Arg Ser Ile Ala Asp Ser Ser Arg Gly Lys Leu Ala Phe Leu  
 530 535 540  
 Phe Thr Gly Gln Gly Ala Gln Thr Leu Gly Met Gly Arg Gly Leu Tyr  
 545 550 555 560  
 Asp Val Trp Ser Ala Phe Arg Glu Ala Phe Asp Leu Cys Val Arg Leu  
 565 570 575  
 Phe Asn Gln Glu Leu Asp Arg Pro Leu Arg Glu Val Met Trp Ala Glu  
 580 585 590  
 Pro Ala Ser Val Asp Ala Ala Leu Leu Asp Gln Thr Ala Phe Thr Gln  
 595 600 605  
 Pro Ala Leu Phe Thr Phe Glu Tyr Ala Leu Ala Ala Leu Trp Arg Ser  
 610 615 620  
 Trp Gly Val Glu Pro Glu Leu Val Ala Gly His Ser Ile Gly Glu Leu  
 625 630 635 640  
 Val Ala Ala Cys Val Ala Gly Val Phe Ser Leu Glu Asp Ala Val Phe  
 645 650 655  
 Leu Val Ala Ala Arg Gly Arg Leu Met Gln Ala Leu Pro Ala Gly Gly  
 660 665 670  
 Ala Met Val Ser Ile Glu Ala Pro Glu Ala Asp Val Ala Ala Ala Val  
 675 680 685  
 Ala Pro His Ala Ala Ser Val Ser Ile Ala Ala Val Asn Ala Pro Asp  
 690 695 700  
 Gln Val Val Ile Ala Gly Ala Gly Gln Pro Val His Ala Ile Ala Ala  
 705 710 715 720  
 Ala Met Ala Ala Arg Gly Ala Arg Thr Lys Ala Leu His Val Ser His  
 725 730 735  
 Ala Phe His Ser Pro Leu Met Ala Pro Met Leu Glu Ala Phe Gly Arg  
 740 745 750



Val Ala Glu Ser Val Ser Tyr Arg Arg Pro Ser Ile Val Leu Val Ser  
 755 760 765  
 Asn Leu Ser Gly Lys Ala Cys Thr Asp Glu Val Ser Ser Pro Gly Tyr  
 770 775 780  
 Trp Val Arg His Ala Arg Glu Val Val Arg Phe Ala Asp Gly Val Lys  
 785 790 795 800  
 Ala Leu His Ala Ala Gly Ala Gly Thr Phe Val Glu Val Gly Pro Lys  
 805 810 815  
 Ser Thr Leu Leu Gly Leu Val Pro Ala Cys Met Pro Asp Ala Arg Pro  
 820 825 830  
 Ala Leu Leu Ala Ser Ser Arg Ala Gly Arg Asp Glu Pro Ala Thr Val  
 835 840 845  
 Leu Glu Ala Leu Gly Gly Leu Trp Ala Val Gly Gly Leu Val Ser Trp  
 850 855 860  
 Ala Gly Leu Phe Pro Ser Gly Gly Arg Arg Val Pro Leu Pro Thr Tyr  
 865 870 875 880  
 Pro Trp Gln Arg Glu Arg Tyr Trp Ile Asp Thr Lys Ala Asp Asp Ala  
 885 890 895  
 Ala Arg Gly Asp Arg Arg Ala Pro Gly Ala Gly His Asp Glu Val Glu  
 900 905 910  
 Glu Gly Gly Ala Val Arg Gly Gly Asp Arg Arg Ser Ala Arg Leu Asp  
 915 920 925  
 His Pro Pro Pro Glu Ser Gly Arg Arg Glu Lys Val Glu Ala Ala Gly  
 930 935 940  
 Asp Arg Pro Phe Arg Leu Glu Ile Asp Glu Pro Gly Val Leu Asp His  
 945 950 955 960  
 Leu Val Leu Arg Val Thr Glu Arg Arg Ala Pro Gly Leu Gly Glu Val  
 965 970 975  
 Glu Ile Ala Val Asp Ala Ala Gly Leu Ser Phe Asn Asp Val Gln Leu  
 980 985 990  
 Ala Leu Gly Met Val Pro Asp Asp Leu Pro Gly Lys Pro Asn Pro Pro  
 995 1000 1005  
 Leu Leu Leu Gly Gly Glu Cys Ala Gly Arg Ile Val Ala Val Gly Glu  
 1010 1015 1020  
 Gly Val Asn Gly Leu Val Val Gly Gln Pro Val Ile Ala Leu Ser Ala  
 1025 1030 1035 1040  
 Gly Ala Phe Ala Thr His Val Thr Thr Ser Ala Ala Leu Val Leu Pro  
 1045 1050 1055  
 Arg Pro Gln Ala Leu Ser Ala Ile Glu Ala Ala Ala Met Pro Val Ala  
 1060 1065 1070  
 Tyr Leu Thr Ala Trp Tyr Ala Leu Asp Arg Ile Ala Arg Leu Gln Pro  
 1075 1080 1085  
 Gly Glu Arg Val Leu Ile His Ala Ala Thr Gly Gly Val Gly Leu Ala  
 1090 1095 1100

- 22 -

Ala Val Gln Trp Ala Gln His Val Gly Ala Glu Val His Ala Thr Ala  
 1105 1110 1115 1120

Gly Thr Pro Glu Lys Arg Ala Tyr Leu Glu Ser Leu Gly Val Arg Tyr  
 1125 1130 1135

Val Ser Asp Ser Arg Ser Asp Arg Phe Val Ala Asp Val Arg Ala Trp  
 1140 1145 1150

Thr Gly Gly Glu Gly Val Asp Val Val Leu Asn Ser Leu Ser Gly Glu  
 1155 1160 1165

Leu Ile Asp Lys Ser Phe Asn Leu Leu Arg Ser His Gly Arg Phe Val  
 1170 1175 1180

Glu Leu Gly Lys Arg Asp Cys Tyr Ala Asp Asn Gln Leu Gly Leu Arg  
 1185 1190 1195 1200

Pro Phe Leu Arg Asn Leu Ser Phe Ser Leu Val Asp Leu Arg Gly Met  
 1205 1210 1215

Met Leu Glu Arg Pro Ala Arg Val Arg Ala Leu Leu Glu Glu Leu Leu  
 1220 1225 1230

Gly Leu Ile Ala Ala Gly Val Phe Thr Pro Pro Pro Ile Ala Thr Leu  
 1235 1240 1245

Pro Ile Ala Arg Val Ala Asp Ala Phe Arg Ser Met Ala Gln Ala Gln  
 1250 1255 1260

His Leu Gly Lys Leu Val Leu Thr Leu Gly Asp Pro Glu Val Gln Ile  
 1265 1270 1275 1280

Arg Ile Pro Thr His Ala Gly Ala Gly Pro Ser Thr Gly Asp Arg Asp  
 1285 1290 1295

Leu Leu Asp Arg Leu Ala Ser Ala Ala Pro Ala Ala Arg Ala Ala Ala  
 1300 1305 1310

Leu Glu Ala Phe Leu Arg Thr Gln Val Ser Gln Val Leu Arg Thr Pro  
 1315 1320 1325

Glu Ile Lys Val Gly Ala Glu Ala Leu Phe Thr Arg Leu Gly Met Asp  
 1330 1335 1340

Ser Leu Met Ala Val Glu Leu Arg Asn Arg Ile Glu Ala Ser Leu Lys  
 1345 1350 1355 1360

Leu Lys Leu Ser Thr Thr Phe Leu Ser Thr Ser Pro Asn Ile Ala Leu  
 1365 1370 1375

Leu Ala Gln Asn Leu Leu Asp Ala Leu Ala Thr Ala Leu Ser Leu Glu  
 1380 1385 1390

Arg Val Ala Ala Glu Asn Leu Arg Ala Gly Val Gln Asn Asp Phe Val  
 1395 1400 1405

Ser Ser Gly Ala Asp Gln Asp Trp Glu Ile Ile Ala Leu  
 1410 1415 1420

&lt;210&gt; 3

&lt;211&gt; 1410

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 3

Met Thr Ile Asn Gln Leu Leu Asn Glu Leu Glu His Gln Gly Ile Lys  
 1 5 10 15  
 Leu Ala Ala Asp Gly Glu Arg Leu Gln Ile Gln Ala Pro Lys Asn Ala  
 20 25 30  
 Leu Asn Pro Asn Leu Leu Ala Arg Ile Ser Glu His Lys Ser Thr Ile  
 35 40 45  
 Leu Thr Met Leu Arg Gln Arg Leu Pro Ala Glu Ser Ile Val Pro Ala  
 50 55 60  
 Pro Ala Glu Arg His Ala Pro Phe Pro Leu Thr Asp Ile Gln Glu Ser  
 65 70 75 80  
 Tyr Trp Leu Gly Arg Thr Gly Ala Phe Thr Val Pro Ser Gly Ile His  
 85 90 95  
 Ala Tyr Arg Glu Tyr Asp Cys Thr Asp Leu Asp Val Pro Arg Leu Ser  
 100 105 110  
 Arg Ala Phe Arg Lys Val Val Ala Arg His Asp Met Leu Arg Ala His  
 115 120 125  
 Thr Leu Pro Asp Met Met Gln Val Ile Glu Pro Lys Val Asp Ala Asp  
 130 135 140  
 Ile Glu Ile Ile Asp Leu Arg Gly Leu Asp Arg Ser Thr Arg Glu Ala  
 145 150 155 160  
 Arg Leu Val Ser Leu Arg Asp Ala Met Ser His Arg Ile Tyr Asp Thr  
 165 170 175  
 Glu Arg Pro Pro Leu Tyr His Val Val Ala Val Arg Leu Asp Glu Arg  
 180 185 190  
 Gln Thr Arg Leu Val Leu Ser Ile Asp Leu Ile Asn Val Asp Leu Gly  
 195 200 205  
 Ser Leu Ser Ile Ile Phe Lys Asp Trp Leu Ser Phe Tyr Glu Asp Pro  
 210 215 220  
 Glu Thr Ser Leu Pro Val Leu Glu Leu Ser Tyr Arg Asp Tyr Val Leu  
 225 230 235 240  
 Ala Leu Glu Ser Arg Lys Lys Ser Glu Ala His Gln Arg Ser Met Asp  
 245 250 255  
 Tyr Trp Lys Arg Arg Ile Ala Glu Leu Pro Pro Pro Pro Thr Leu Pro  
 260 265 270  
 Met Lys Ala Asp Pro Ser Thr Leu Lys Glu Ile Arg Phe Arg His Thr  
 275 280 285  
 Glu Gln Trp Leu Pro Ser Asp Ser Trp Gly Arg Leu Lys Arg Arg Val  
 290 295 300  
 Gly Glu Arg Gly Leu Thr Pro Thr Gly Val Ile Leu Ala Ala Phe Ser  
 305 310 315 320  
 Glu Val Ile Gly Arg Trp Ser Ala Ser Pro Arg Phe Thr Leu Asn Ile  
 325 330 335  
 Thr Leu Phe Asn Arg Leu Pro Val His Pro Arg Val Asn Asp Ile Thr

340	345	350
Gly Asp Phe Thr Ser Met Val	Leu Leu Asp Ile Asp Thr Thr Arg Asp	
355	360	365
Lys Ser Phe Glu Gln Arg Ala Lys Arg Ile Gln Glu Gln Leu Trp Glu		
370	375	380
Ala Met Asp His Cys Asp Val Ser Gly Ile Glu Val Gln Arg Glu Ala		
385	390	395
Ala Arg Val Leu Gly Ile Gln Arg Gly Ala Leu Phe Pro Val Val Leu		
405	410	415
Thr Ser Ala Leu Asn Gln Gln Val Val Gly Val Thr Ser Leu Gln Arg		
420	425	430
Leu Gly Thr Pro Val Tyr Thr Ser Thr Gln Thr Pro Gln Leu Leu Leu		
435	440	445
Asp His Gln Leu Tyr Glu His Asp Gly Asp Leu Val Leu Ala Trp Asp		
450	455	460
Ile Val Asp Gly Val Phe Pro Pro Asp Leu Leu Asp Asp Met Leu Glu		
465	470	475
Ala Tyr Val Val Phe Leu Arg Arg Leu Thr Glu Glu Pro Trp Gly Glu		
485	490	495
Gln Val Arg Cys Ser Leu Pro Pro Ala Gln Leu Glu Ala Arg Ala Ser		
500	505	510
Ala Asn Ala Thr Asn Ala Leu Leu Ser Glu His Thr Leu His Gly Leu		
515	520	525
Phe Ala Ala Arg Val Glu Gln Leu Pro Met Gln Leu Ala Val Val Ser		
530	535	540
Ala Arg Lys Thr Leu Thr Tyr Glu Glu Leu Ser Arg Arg Ser Arg Arg		
545	550	555
Leu Gly Ala Arg Leu Arg Glu Gln Gly Ala Arg Pro Asn Thr Leu Val		
565	570	575
Ala Val Val Met Glu Lys Gly Trp Glu Gln Val Val Ala Val Leu Ala		
580	585	590
Val Leu Glu Ser Gly Ala Ala Tyr Val Pro Ile Asp Ala Asp Leu Pro		
595	600	605
Ala Glu Arg Ile His Tyr Leu Leu Asp His Gly Glu Val Lys Leu Val		
610	615	620
Leu Thr Gln Pro Trp Leu Asp Gly Lys Leu Ser Trp Pro Pro Gly Ile		
625	630	635
Gln Arg Leu Leu Val Ser Glu Ala Gly Val Glu Gly Asp Gly Asp Gln		
645	650	655
Pro Pro Met Met Pro Ile Gln Thr Pro Ser Asp Leu Ala Tyr Val Ile		
660	665	670
Tyr Thr Ser Gly Ser Thr Gly Leu Pro Lys Gly Val Met Ile Asp His		
675	680	685
Arg Gly Ala Val Asn Thr Ile Leu Asp Ile Asn Glu Arg Phe Glu Ile		

690					695					700					
Gly	Pro	Gly	Asp	Arg	Val	Leu	Ala	Leu	Ser	Ser	Leu	Ser	Phe	Asp	Leu
705					710					715					720
Ser	Val	Tyr	Asp	Val	Phe	Gly	Ile	Leu	Ala	Ala	Gly	Gly	Thr	Ile	Val
				725					730					735	
Val	Pro	Asp	Ala	Ser	Lys	Leu	Arg	Asp	Pro	Ala	His	Trp	Ala	Glu	Leu
			740					745					750		
Ile	Glu	Arg	Glu	Lys	Val	Thr	Val	Trp	Asn	Ser	Val	Pro	Ala	Leu	Met
		755					760					765			
Arg	Met	Leu	Val	Glu	His	Phe	Glu	Gly	Arg	Pro	Asp	Ser	Leu	Ala	Arg
	770					775					780				
Ser	Leu	Arg	Leu	Ser	Leu	Leu	Ser	Gly	Asp	Trp	Ile	Pro	Val	Gly	Leu
785					790					795					800
Pro	Gly	Glu	Leu	Gln	Ala	Ile	Arg	Pro	Gly	Val	Ser	Val	Ile	Ser	Leu
				805					810					815	
Gly	Gly	Ala	Thr	Glu	Ala	Ser	Ile	Trp	Ser	Ile	Gly	Tyr	Pro	Val	Arg
			820					825					830		
Asn	Val	Asp	Leu	Ser	Trp	Ala	Ser	Ile	Pro	Tyr	Gly	Arg	Pro	Leu	Arg
		835					840					845			
Asn	Gln	Thr	Phe	His	Val	Leu	Asp	Glu	Ala	Leu	Glu	Pro	Arg	Pro	Val
	850					855					860				
Trp	Val	Pro	Gly	Gln	Leu	Tyr	Ile	Gly	Gly	Val	Gly	Leu	Ala	Leu	Gly
865				870					875						880
Tyr	Trp	Arg	Asp	Glu	Glu	Lys	Thr	Arg	Lys	Ser	Phe	Leu	Val	His	Pro
				885					890					895	
Glu	Thr	Gly	Glu	Arg	Leu	Tyr	Lys	Thr	Gly	Asp	Leu	Gly	Arg	Tyr	Leu
			900					905					910		
Pro	Asp	Gly	Asn	Ile	Glu	Phe	Met	Gly	Arg	Glu	Asp	Asn	Gln	Ile	Lys
		915					920					925			
Leu	Arg	Gly	Tyr	Arg	Val	Glu	Leu	Gly	Glu	Ile	Glu	Glu	Thr	Leu	Lys
	930					935					940				
Ser	His	Pro	Asn	Val	Arg	Asp	Ala	Val	Ile	Val	Pro	Val	Gly	Asn	Asp
945					950					955				960	
Ala	Ala	Asn	Lys	Leu	Leu	Leu	Ala	Tyr	Val	Val	Pro	Glu	Gly	Thr	Arg
			965					970						975	
Arg	Arg	Ala	Ala	Glu	Gln	Asp	Ala	Ser	Leu	Lys	Thr	Glu	Arg	Ile	Asp
			980					985					990		
Ala	Arg	Ala	His	Ala	Ala	Glu	Ala	Asp	Gly	Leu	Ser	Asp	Gly	Glu	Arg
		995					1000					1005			
Val	Gln	Phe	Lys	Leu	Ala	Arg	His	Gly	Leu	Arg	Arg	Asp	Leu	Asp	Gly
	1010					1015					1020				
Lys	Pro	Val	Val	Asp	Leu	Thr	Gly	Gln	Asp	Pro	Arg	Glu	Ala	Gly	Leu
1025					1030				1035					1040	
Asp	Val	Tyr	Ala	Arg	Arg	Arg	Ser	Val	Arg	Thr	Phe	Leu	Glu	Ala	Pro

1045	1050	1055
Ile Pro Phe Val Glu Phe Gly Arg Phe Leu Ser Cys Leu Ser Ser Val 1060	1065	1070
Glu Pro Asp Gly Ala Thr Leu Pro Lys Phe Arg Tyr Pro Ser Ala Gly 1075	1080	1085
Ser Thr Tyr Pro Val Gln Thr Tyr Ala Tyr Val Lys Ser Gly Arg Ile 1090	1095	1100
Glu Gly Val Asp Glu Gly Phe Tyr Tyr Tyr His Pro Phe Glu His Arg 1105	1110	1115
Leu Leu Lys Leu Ser Asp His Gly Ile Glu Arg Gly Ala His Val Arg 1125	1130	1135
Gln Asn Phe Asp Val Phe Asp Glu Ala Ala Phe Asn Leu Leu Phe Val 1140	1145	1150
Gly Arg Ile Asp Ala Ile Glu Ser Leu Tyr Gly Ser Ser Ser Arg Glu 1155	1160	1165
Phe Cys Leu Leu Glu Ala Gly Tyr Met Ala Gln Leu Leu Met Glu Gln 1170	1175	1180
Ala Pro Ser Cys Asn Ile Gly Val Cys Pro Val Gly Gln Phe Asn Phe 1185	1190	1195
Glu Gln Val Arg Pro Val Leu Asp Leu Arg His Ser Asp Val Tyr Val 1205	1210	1215
His Gly Met Leu Gly Gly Arg Val Asp Pro Arg Gln Phe Gln Val Cys 1220	1225	1230
Thr Leu Gly Gln Asp Ser Ser Pro Arg Arg Ala Thr Thr Arg Gly Ala 1235	1240	1245
Pro Pro Gly Arg Glu Gln His Phe Ala Asp Met Leu Arg Asp Phe Leu 1250	1255	1260
Arg Thr Lys Leu Pro Glu Tyr Met Val Pro Thr Val Phe Val Glu Leu 1265	1270	1275
Asp Ala Leu Pro Leu Thr Ser Asn Gly Lys Val Asp Arg Lys Ala Leu 1285	1290	1295
Arg Glu Arg Lys Asp Thr Ser Ser Pro Arg His Ser Gly His Thr Ala 1300	1305	1310
Pro Arg Asp Ala Leu Glu Glu Ile Leu Val Ala Val Val Arg Glu Val 1315	1320	1325
Leu Gly Leu Glu Val Val Gly Leu Gln Gln Ser Phe Val Asp Leu Gly 1330	1335	1340
Ala Thr Ser Ile His Ile Val Arg Met Arg Ser Leu Leu Gln Lys Arg 1345	1350	1355
Leu Asp Arg Glu Ile Ala Ile Thr Glu Leu Phe Gln Tyr Pro Asn Leu 1365	1370	1375
Gly Ser Leu Ala Ser Gly Leu Arg Arg Asp Ser Arg Asp Leu Asp Gln 1380	1385	1390
Arg Pro Asn Met Gln Asp Arg Val Glu Val Arg Arg Lys Gly Arg Arg		



1395

1400

1405

Arg Ser  
1410

&lt;210&gt; 4

&lt;211&gt; 1832

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 4

Met Glu Glu Gln Glu Ser Ser Ala Ile Ala Val Ile Gly Met Ser Gly  
1 5 10 15Arg Phe Pro Gly Ala Arg Asp Leu Asp Glu Phe Trp Arg Asn Leu Arg  
20 25 30Asp Gly Thr Glu Ala Val Gln Arg Phe Ser Glu Gln Glu Leu Ala Ala  
35 40 45Ser Gly Val Asp Pro Ala Leu Val Leu Asp Pro Ser Tyr Val Arg Ala  
50 55 60Gly Ser Val Leu Glu Asp Val Asp Arg Phe Asp Ala Ala Phe Phe Gly  
65 70 75 80Ile Ser Pro Arg Glu Ala Glu Leu Met Asp Pro Gln His Arg Ile Phe  
85 90 95Met Glu Cys Ala Trp Glu Ala Leu Glu Asn Ala Gly Tyr Asp Pro Thr  
100 105 110Ala Tyr Glu Gly Ser Ile Gly Val Tyr Ala Gly Ala Asn Met Ser Ser  
115 120 125Tyr Leu Thr Ser Asn Leu His Glu His Pro Ala Met Met Arg Trp Pro  
130 135 140Gly Trp Phe Gln Thr Leu Ile Gly Asn Asp Lys Asp Tyr Leu Ala Thr  
145 150 155 160His Val Ser Tyr Arg Leu Asn Leu Arg Gly Pro Ser Ile Ser Val Gln  
165 170 175Thr Ala Cys Ser Thr Ser Leu Val Ala Val His Leu Ala Cys Met Ser  
180 185 190Leu Leu Asp Arg Glu Cys Asp Met Ala Leu Ala Gly Gly Ile Thr Val  
195 200 205Arg Ile Pro His Arg Ala Gly Tyr Val Tyr Ala Glu Gly Gly Ile Phe  
210 215 220Ser Pro Asp Gly His Cys Arg Ala Phe Asp Ala Lys Ala Asn Gly Thr  
225 230 235 240Ile Met Gly Asn Gly Cys Gly Val Val Leu Leu Lys Pro Leu Asp Arg  
245 250 255Ala Leu Ser Asp Gly Asp Pro Val Arg Ala Val Ile Leu Gly Ser Ala  
260 265 270Thr Asn Asn Asp Gly Ala Arg Lys Ile Gly Phe Thr Ala Pro Ser Glu  
275 280 285

Val Gly Gln Ala Gln Ala Ile Met Glu Ala Leu Ala Leu Ala Gly Val  
 290 295 300  
 Glu Ala Arg Ser Ile Gln Tyr Ile Glu Thr His Gly Thr Gly Thr Leu  
 305 310 315 320  
 Leu Gly Asp Ala Ile Glu Thr Ala Ala Leu Arg Arg Val Phe Gly Arg  
 325 330 335  
 Asp Ala Ser Ala Arg Arg Ser Cys Ala Ile Gly Ser Val Lys Thr Gly  
 340 345 350  
 Ile Gly His Leu Glu Ser Ala Ala Gly Ile Ala Gly Leu Ile Lys Thr  
 355 360 365  
 Val Leu Ala Leu Glu His Arg Gln Leu Pro Pro Ser Leu Asn Phe Glu  
 370 375 380  
 Ser Pro Asn Pro Ser Ile Asp Phe Ala Ser Ser Pro Phe Tyr Val Asn  
 385 390 395 400  
 Thr Ser Leu Lys Asp Trp Asn Thr Gly Ser Thr Pro Arg Arg Ala Gly  
 405 410 415  
 Val Ser Ser Phe Gly Ile Gly Gly Thr Asn Ala His Val Val Leu Glu  
 420 425 430  
 Glu Ala Pro Ala Ala Lys Leu Pro Ala Ala Ala Pro Ala Arg Ser Ala  
 435 440 445  
 Glu Leu Phe Val Val Ser Ala Lys Ser Ala Ala Ala Leu Asp Ala Ala  
 450 455 460  
 Ala Ala Arg Leu Arg Asp His Leu Gln Ala His Gln Gly Ile Ser Leu  
 465 470 475 480  
 Gly Asp Val Ala Phe Ser Leu Ala Thr Thr Arg Ser Pro Met Glu His  
 485 490 495  
 Arg Leu Ala Met Ala Ala Pro Ser Arg Glu Ala Leu Arg Glu Gly Leu  
 500 505 510  
 Asp Ala Ala Ala Arg Gly Gln Thr Pro Pro Gly Ala Val Arg Gly Arg  
 515 520 525  
 Cys Ser Pro Gly Asn Val Pro Lys Val Val Phe Val Phe Pro Gly Gln  
 530 535 540  
 Gly Ser Gln Trp Val Gly Met Gly Arg Gln Leu Leu Ala Glu Glu Pro  
 545 550 555 560  
 Val Phe His Ala Ala Leu Ser Ala Cys Asp Arg Ala Ile Gln Ala Glu  
 565 570 575  
 Ala Gly Trp Ser Leu Leu Ala Glu Leu Ala Ala Asp Glu Gly Ser Ser  
 580 585 590  
 Gln Leu Glu Arg Ile Asp Val Val Gln Pro Val Leu Phe Ala Leu Ala  
 595 600 605  
 Val Ala Phe Ala Ala Leu Trp Arg Ser Trp Gly Val Ala Pro Asp Val  
 610 615 620  
 Val Ile Gly His Ser Met Gly Glu Val Ala Ala His Val Ala Gly  
 625 630 635 640

- 29 -

Ala Leu Ser Leu Glu Asp Ala Val Ala Ile Ile Cys Arg Arg Ser Arg  
 645 650 655  
 Leu Leu Arg Arg Ile Ser Gly Gln Gly Glu Met Ala Val Thr Glu Leu  
 660 665 670  
 Ser Leu Ala Glu Ala Glu Ala Ala Leu Arg Gly Tyr Glu Asp Arg Val  
 675 680 685  
 Ser Val Ala Val Ser Asn Ser Pro Arg Ser Thr Val Leu Ser Gly Glu  
 690 695 700  
 Pro Ala Ala Ile Gly Glu Val Leu Ser Ser Leu Asn Ala Lys Gly Val  
 705 710 715 720  
 Phe Cys Arg Arg Val Lys Val Asp Val Ala Ser His Ser Pro Gln Val  
 725 730 735  
 Asp Pro Leu Arg Glu Asp Leu Leu Ala Ala Leu Gly Gly Leu Arg Pro  
 740 745 750  
 Gly Ala Ala Ala Val Pro Met Arg Ser Thr Val Thr Gly Ala Met Val  
 755 760 765  
 Ala Gly Pro Glu Leu Gly Ala Asn Tyr Trp Met Asn Asn Leu Arg Gln  
 770 775 780  
 Pro Val Arg Phe Ala Glu Val Val Gln Ala Gln Leu Gln Gly Gly His  
 785 790 795 800  
 Gly Leu Phe Val Glu Met Ser Pro His Pro Ile Leu Thr Thr Ser Val  
 805 810 815  
 Glu Glu Met Arg Arg Ala Ala Gln Arg Ala Gly Ala Ala Val Gly Ser  
 820 825 830  
 Leu Arg Arg Gly Gln Asp Glu Arg Pro Ala Met Leu Glu Ala Leu Gly  
 835 840 845  
 Thr Leu Trp Ala Gln Gly Tyr Pro Val Pro Trp Gly Arg Leu Phe Pro  
 850 855 860  
 Ala Gly Gly Arg Arg Val Pro Leu Pro Thr Tyr Pro Trp Gln Arg Glu  
 865 870 875 880  
 Arg Tyr Trp Ile Glu Ala Pro Ala Lys Ser Ala Ala Gly Asp Arg Arg  
 885 890 895  
 Gly Val Arg Ala Gly Gly His Pro Leu Leu Gly Glu Met Gln Thr Leu  
 900 905 910  
 Ser Thr Gln Thr Ser Thr Arg Leu Trp Glu Thr Thr Leu Asp Leu Lys  
 915 920 925  
 Arg Leu Pro Trp Leu Gly Asp His Arg Val Gln Gly Ala Val Val Phe  
 930 935 940  
 Pro Gly Ala Ala Tyr Leu Glu Met Ala Ile Ser Ser Gly Ala Glu Ala  
 945 950 955 960  
 Leu Gly Asp Gly Pro Leu Gln Ile Thr Asp Val Val Leu Ala Glu Ala  
 965 970 975  
 Leu Ala Phe Ala Gly Asp Ala Ala Val Leu Val Gln Val Val Thr Thr  
 980 985 990

Glu Gln Pro Ser Gly Arg Leu Gln Phe Gln Ile Ala Ser Arg Ala Pro  
 995 1000 1005  
 Gly Ala Gly His Ala Ser Phe Arg Val His Ala Arg Gly Ala Leu Leu  
 1010 1015 1020  
 Arg Val Glu Arg Thr Glu Val Pro Ala Gly Leu Thr Leu Ser Ala Val  
 1025 1030 1035 1040  
 Arg Ala Arg Leu Gln Ala Ser Ile Pro Ala Ala Ala Thr Tyr Ala Glu  
 1045 1050 1055  
 Leu Thr Glu Met Gly Leu Gln Tyr Gly Pro Ala Phe Gln Gly Ile Ala  
 1060 1065 1070  
 Glu Leu Trp Arg Gly Glu Gly Glu Ala Leu Gly Arg Val Arg Leu Pro  
 1075 1080 1085  
 Asp Ala Ala Gly Ser Ala Ala Glu Tyr Arg Leu His Pro Ala Leu Leu  
 1090 1095 1100  
 Asp Ala Cys Phe Gln Ile Val Gly Ser Leu Phe Ala Arg Ser Gly Glu  
 1105 1110 1115 1120  
 Ala Thr Pro Trp Val Pro Val Glu Leu Gly Ser Leu Arg Leu Leu Gln  
 1125 1130 1135  
 Arg Pro Ser Gly Glu Leu Trp Cys His Ala Arg Val Val Asn His Gly  
 1140 1145 1150  
 His Gln Thr Pro Asp Arg Gln Gly Ala Asp Phe Trp Val Val Asp Ser  
 1155 1160 1165  
 Ser Gly Ala Val Val Ala Glu Val Cys Gly Leu Val Ala Gln Arg Leu  
 1170 1175 1180  
 Pro Gly Gly Val Arg Arg Arg Glu Glu Asp Asp Trp Phe Leu Glu Leu  
 1185 1190 1195 1200  
 Glu Trp Glu Pro Ala Ala Val Gly Thr Ala Lys Val Asn Ala Gly Arg  
 1205 1210 1215  
 Trp Leu Leu Leu Gly Gly Gly Gly Gly Leu Gly Ala Ala Leu Arg Ala  
 1220 1225 1230  
 Met Leu Glu Ala Gly Gly His Ala Val Val His Ala Ala Glu Asn Asn  
 1235 1240 1245  
 Thr Ser Ala Ala Gly Val Arg Ala Leu Leu Ala Lys Ala Phe Asp Gly  
 1250 1255 1260  
 Gln Ala Pro Thr Ala Val Val His Leu Gly Ser Leu Asp Gly Gly Gly  
 1265 1270 1275 1280  
 Glu Leu Asp Pro Gly Leu Gly Ala Gln Gly Ala Leu Asp Ala Pro Arg  
 1285 1290 1295  
 Ser Ala Asp Val Ser Pro Asp Ala Leu Asp Pro Ala Leu Val Arg Gly  
 1300 1305 1310  
 Cys Asp Ser Val Leu Trp Thr Val Gln Ala Leu Ala Gly Met Gly Phe  
 1315 1320 1325  
 Arg Asp Ala Pro Arg Leu Trp Leu Leu Thr Arg Gly Ala Gln Ala Val  
 1330 1335 1340

Gly Ala Gly Asp Val Ser Val Thr Gln Ala Pro Leu Leu Gly Leu Gly  
 1345 1350 1355 1360  
 Arg Val Ile Ala Met Glu His Ala Asp Leu Arg Cys Ala Arg Val Asp  
 1365 1370 1375  
 Leu Asp Pro Ala Arg Pro Glu Gly Glu Leu Ala Ala Leu Leu Ala Glu  
 1380 1385 1390  
 Leu Leu Ala Asp Asp Ala Glu Ala Glu Val Ala Leu Arg Gly Gly Glu  
 1395 1400 1405  
 Arg Cys Val Ala Arg Ile Val Arg Arg Gln Pro Glu Thr Arg Pro Arg  
 1410 1415 1420  
 Gly Arg Ile Glu Ser Cys Val Pro Thr Asp Val Thr Ile Arg Ala Asp  
 1425 1430 1435 1440  
 Ser Thr Tyr Leu Val Thr Gly Gly Leu Gly Gly Leu Gly Leu Ser Val  
 1445 1450 1455  
 Ala Gly Trp Leu Ala Glu Arg Gly Ala Gly His Leu Val Leu Val Gly  
 1460 1465 1470  
 Arg Ser Gly Ala Ala Ser Val Glu Gln Arg Ala Ala Val Ala Ala Leu  
 1475 1480 1485  
 Glu Ala Arg Gly Ala Arg Val Thr Val Ala Lys Ala Asp Val Ala Asp  
 1490 1495 1500  
 Arg Ala Gln Leu Glu Arg Ile Leu Arg Glu Val Thr Thr Ser Gly Met  
 1505 1510 1515 1520  
 Pro Leu Arg Gly Val Val His Ala Ala Gly Ile Leu Asp Asp Gly Leu  
 1525 1530 1535  
 Leu Met Gln Gln Thr Pro Ala Arg Phe Arg Lys Val Met Ala Pro Lys  
 1540 1545 1550  
 Val Gln Gly Ala Leu His Leu His Ala Leu Thr Arg Glu Ala Pro Leu  
 1555 1560 1565  
 Ser Phe Phe Val Leu Tyr Ala Ser Gly Val Gly Leu Leu Gly Ser Pro  
 1570 1575 1580  
 Gly Gln Gly Asn Tyr Ala Ala Ala Asn Thr Phe Leu Asp Ala Leu Ala  
 1585 1590 1595 1600  
 His His Arg Arg Ala Gln Gly Leu Pro Ala Leu Ser Val Asp Trp Gly  
 1605 1610 1615  
 Leu Phe Ala Glu Val Gly Met Ala Ala Ala Gln Glu Asp Arg Gly Ala  
 1620 1625 1630  
 Arg Leu Val Ser Arg Gly Met Arg Ser Leu Thr Pro Asp Glu Gly Leu  
 1635 1640 1645  
 Ser Ala Leu Ala Arg Leu Leu Glu Ser Gly Arg Ala Gln Val Gly Val  
 1650 1655 1660  
 Met Pro Val Asn Pro Arg Leu Trp Val Glu Leu Tyr Pro Ala Ala Ala  
 1665 1670 1675 1680  
 Ser Ser Arg Met Leu Ser Arg Leu Val Thr Ala His Arg Ala Ser Ala  
 1685 1690 1695

- 32 -

Gly Gly Pro Ala Gly Asp Gly Asp Leu Arg Arg Leu Ala Ala Ala  
 1700 1705 1710

Glu Pro Ser Ala Arg Ser Ala Leu Leu Glu Pro Leu Leu Arg Ala Gln  
 1715 1720 1725

Ile Ser Gln Val Leu Arg Leu Pro Glu Gly Lys Ile Glu Val Asp Ala  
 1730 1735 1740

Pro Leu Thr Ser Leu Gly Met Asn Ser Leu Met Gly Leu Glu Leu Arg  
 1745 1750 1755 1760

Asn Arg Ile Glu Ala Met Leu Gly Ile Thr Val Pro Ala Thr Leu Leu  
 1765 1770 1775

Trp Thr Tyr Pro Thr Val Ala Ala Leu Ser Gly His Leu Ala Arg Glu  
 1780 1785 1790

Ala Cys Glu Ala Ala Pro Val Glu Ser Pro His Thr Thr Ala Asp Ser  
 1795 1800 1805

Ala Val Glu Ile Glu Glu Met Ser Gln Asp Asp Leu Thr Gln Leu Ile  
 1810 1815 1820

Ala Ala Lys Phe Lys Ala Leu Thr  
 1825 1830

<210> 5

<211> 7257

<212> PRT

<213> Sorangium cellulosum

<400> 5

Met Thr Thr Arg Gly Pro Thr Ala Gln Gln Asn Pro Leu Lys Gln Ala  
 1 5 10 15

Ala Ile Ile Ile Gln Arg Leu Glu Glu Arg Leu Ala Gly Leu Ala Gln  
 20 25 30

Ala Glu Leu Glu Arg Thr Glu Pro Ile Ala Ile Val Gly Ile Gly Cys  
 35 40 45

Arg Phe Pro Gly Gly Ala Asp Ala Pro Glu Ala Phe Trp Glu Leu Leu  
 50 55 60

Asp Ala Glu Arg Asp Ala Val Gln Pro Leu Asp Met Arg Trp Ala Leu  
 65 70 75 80

Val Gly Val Ala Pro Val Glu Ala Val Pro His Trp Ala Gly Leu Leu  
 85 90 95

Thr Glu Pro Ile Asp Cys Phe Asp Ala Ala Phe Phe Gly Ile Ser Pro  
 100 105 110

Arg Glu Ala Arg Ser Leu Asp Pro Gln His Arg Leu Leu Leu Glu Val  
 115 120 125

Ala Trp Glu Gly Leu Glu Asp Ala Gly Ile Pro Pro Arg Ser Ile Asp  
 130 135 140

Gly Ser Arg Thr Gly Val Phe Val Gly Ala Phe Thr Ala Asp Tyr Ala  
 145 150 155 160

Arg Thr Val Ala Arg Leu Pro Arg Glu Glu Arg Asp Ala Tyr Ser Ala  
 165 170 175



Thr Gly Asn Met Leu Ser Ile Ala Ala Gly Arg Leu Ser Tyr Thr Leu  
 180 185 190  
 Gly Leu Gln Gly Pro Cys Leu Thr Val Asp Thr Ala Cys Ser Ser Ser  
 195 200 205  
 Leu Val Ala Ile His Leu Ala Cys Arg Ser Leu Arg Ala Gly Glu Ser  
 210 215 220  
 Asp Leu Ala Leu Ala Gly Gly Val Ser Ala Leu Leu Ser Pro Asp Met  
 225 230 235 240  
 Met Glu Ala Ala Ala Arg Thr Gln Ala Leu Ser Pro Asp Gly Arg Cys  
 245 250 255  
 Arg Thr Phe Asp Ala Ser Ala Asn Gly Phe Val Arg Gly Glu Gly Cys  
 260 265 270  
 Gly Leu Val Val Leu Lys Arg Leu Ser Asp Ala Gln Arg Asp Gly Asp  
 275 280 285  
 Arg Ile Trp Ala Leu Ile Arg Gly Ser Ala Ile Asn His Asp Gly Arg  
 290 295 300  
 Ser Thr Gly Leu Thr Ala Pro Asn Val Leu Ala Gln Glu Thr Val Leu  
 305 310 315 320  
 Arg Glu Ala Leu Arg Ser Ala His Val Glu Ala Gly Ala Val Asp Tyr  
 325 330 335  
 Val Glu Thr His Gly Thr Gly Thr Ser Leu Gly Asp Pro Ile Glu Val  
 340 345 350  
 Glu Ala Leu Arg Ala Thr Val Gly Pro Ala Arg Ser Asp Gly Thr Arg  
 355 360 365  
 Cys Val Leu Gly Ala Val Lys Thr Asn Ile Gly His Leu Glu Ala Ala  
 370 375 380  
 Ala Gly Val Ala Gly Leu Ile Lys Ala Ala Leu Ser Leu Thr His Glu  
 385 390 395 400  
 Arg Ile Pro Arg Asn Leu Asn Phe Arg Thr Leu Asn Pro Arg Ile Arg  
 405 410 415  
 Leu Glu Gly Ser Ala Leu Ala Leu Ala Thr Glu Pro Val Pro Trp Pro  
 420 425 430  
 Arg Thr Asp Arg Pro Arg Phe Ala Gly Val Ser Ser Phe Gly Met Ser  
 435 440 445  
 Gly Thr Asn Ala His Val Val Leu Glu Glu Ala Pro Ala Val Glu Leu  
 450 455 460  
 Trp Pro Ala Ala Pro Glu Arg Ser Ala Glu Leu Leu Val Leu Ser Gly  
 465 470 475 480  
 Lys Ser Glu Gly Ala Leu Asp Ala Gln Ala Ala Arg Leu Arg Glu His  
 485 490 495  
 Leu Asp Met His Pro Glu Leu Gly Leu Gly Asp Val Ala Phe Ser Leu  
 500 505 510  
 Ala Thr Thr Arg Ser Ala Met Ser His Arg Leu Ala Val Ala Val Thr  
 515 520 525

Ser Arg Glu Gly Leu Leu Ala Ala Leu Ser Ala Val Ala Gln Gly Gln  
 530 535 540  
 Thr Pro Ala Gly Ala Ala Arg Cys Ile Ala Ser Ser Ser Arg Gly Lys  
 545 550 555 560  
 Leu Ala Phe Leu Phe Thr Gly Gln Gly Ala Gln Thr Pro Gly Met Gly  
 565 570 575  
 Arg Gly Leu Cys Ala Ala Trp Pro Ala Phe Arg Glu Ala Phe Asp Arg  
 580 585 590  
 Cys Val Ala Leu Phe Asp Arg Glu Leu Asp Arg Pro Leu Arg Glu Val  
 595 600 605  
 Met Trp Ala Glu Ala Gly Ser Ala Glu Ser Leu Leu Leu Asp Gln Thr  
 610 615 620  
 Ala Phe Thr Gln Pro Ala Leu Phe Ala Val Glu Tyr Ala Leu Thr Ala  
 625 630 635 640  
 Leu Trp Arg Ser Trp Gly Val Glu Pro Glu Leu Leu Val Gly His Ser  
 645 650 655  
 Ile Gly Glu Leu Val Ala Ala Cys Val Ala Gly Val Phe Ser Leu Glu  
 660 665 670  
 Asp Gly Val Arg Leu Val Ala Ala Arg Gly Arg Leu Met Gln Gly Leu  
 675 680 685  
 Ser Ala Gly Gly Ala Met Val Ser Leu Gly Ala Pro Glu Ala Glu Val  
 690 695 700  
 Ala Ala Ala Val Ala Pro His Ala Ala Ser Val Ser Ile Ala Ala Val  
 705 710 715 720  
 Asn Gly Pro Glu Gln Val Val Ile Ala Gly Val Glu Gln Ala Val Gln  
 725 730 735  
 Ala Ile Ala Ala Gly Phe Ala Ala Arg Gly Ala Arg Thr Lys Arg Leu  
 740 745 750  
 His Val Ser His Ala Phe His Ser Pro Leu Met Glu Pro Met Leu Glu  
 755 760 765  
 Glu Phe Gly Arg Val Ala Ala Ser Val Thr Tyr Arg Arg Pro Ser Val  
 770 775 780  
 Ser Leu Val Ser Asn Leu Ser Gly Lys Val Val Thr Asp Glu Leu Ser  
 785 790 795 800  
 Ala Pro Gly Tyr Trp Val Arg His Val Arg Glu Ala Val Arg Phe Ala  
 805 810 815  
 Asp Gly Val Lys Ala Leu His Glu Ala Gly Ala Gly Thr Phe Val Glu  
 820 825 830  
 Val Gly Pro Lys Pro Thr Leu Leu Gly Leu Leu Pro Ala Cys Leu Pro  
 835 840 845  
 Glu Ala Glu Pro Thr Leu Leu Ala Ser Leu Arg Ala Gly Arg Glu Glu  
 850 855 860  
 Ala Ala Gly Val Leu Glu Ala Leu Gly Arg Leu Trp Ala Ala Gly Gly  
 865 870 875 880

Ser Val Ser Trp Pro Gly Val Phe Pro Thr Ala Gly Arg Arg Val Pro  
 885 890 895  
 Leu Pro Thr Tyr Pro Trp Gln Arg Gln Arg Tyr Trp Ile Glu Ala Pro  
 900 905 910  
 Ala Glu Gly Leu Gly Ala Thr Ala Ala Asp Ala Leu Ala Gln Trp Phe  
 915 920 925  
 Tyr Arg Val Asp Trp Pro Glu Met Pro Arg Ser Ser Val Asp Ser Arg  
 930 935 940  
 Arg Ala Arg Ser Gly Gly Trp Leu Val Leu Ala Asp Arg Gly Gly Val  
 945 950 955 960  
 Gly Glu Ala Ala Ala Ala Ala Leu Ser Ser Gln Gly Cys Ser Cys Ala  
 965 970 975  
 Val Leu His Ala Pro Ala Glu Ala Ser Ala Val Ala Glu Gln Val Thr  
 980 985 990  
 Gln Ala Leu Gly Gly Arg Asn Asp Trp Gln Gly Val Leu Tyr Leu Trp  
 995 1000 1005  
 Gly Leu Asp Ala Val Val Glu Ala Gly Ala Ser Ala Glu Glu Val Ala  
 1010 1015 1020  
 Lys Val Thr His Leu Ala Ala Ala Pro Val Leu Ala Leu Ile Gln Ala  
 1025 1030 1035 1040  
 Leu Gly Thr Gly Pro Arg Ser Pro Arg Leu Trp Ile Val Thr Arg Gly  
 1045 1050 1055  
 Ala Cys Thr Val Gly Gly Glu Pro Asp Ala Ala Pro Cys Gln Ala Ala  
 1060 1065 1070  
 Leu Trp Gly Met Gly Arg Val Ala Ala Leu Glu His Pro Gly Ser Trp  
 1075 1080 1085  
 Gly Gly Leu Val Asp Leu Asp Pro Glu Glu Ser Pro Thr Glu Val Glu  
 1090 1095 1100  
 Ala Leu Val Ala Glu Leu Leu Ser Pro Asp Ala Glu Asp Gln Leu Ala  
 1105 1110 1115 1120  
 Phe Arg Gln Gly Arg Arg Arg Ala Ala Arg Leu Val Ala Ala Pro Pro  
 1125 1130 1135  
 Glu Gly Asn Ala Ala Pro Val Ser Leu Ser Ala Glu Gly Ser Tyr Leu  
 1140 1145 1150  
 Val Thr Gly Gly Leu Gly Ala Leu Gly Leu Leu Val Ala Arg Trp Leu  
 1155 1160 1165  
 Val Glu Arg Gly Ala Gly His Leu Val Leu Ile Ser Arg His Gly Leu  
 1170 1175 1180  
 Pro Asp Arg Glu Glu Trp Gly Arg Asp Gln Pro Pro Glu Val Arg Ala  
 1185 1190 1195 1200  
 Arg Ile Ala Ala Ile Glu Ala Leu Glu Ala Gln Gly Ala Arg Val Thr  
 1205 1210 1215  
 Val Ala Ala Val Asp Val Ala Asp Ala Glu Gly Met Ala Ala Leu Leu  
 1220 1225 1230

Ala Ala Val Glu Pro Pro Leu Arg Gly Val Val His Ala Ala Gly Leu  
 1235 1240 1245  
 Leu Asp Asp Gly Leu Leu Ala His Gln Asp Ala Gly Arg Leu Ala Arg  
 1250 1255 1260  
 Val Leu Arg Pro Lys Val Glu Gly Ala Trp Val Leu His Thr Leu Thr  
 1265 1270 1275 1280  
 Arg Glu Gln Pro Leu Asp Leu Phe Val Leu Phe Ser Ser Ala Ser Gly  
 1285 1290 1295  
 Val Phe Gly Ser Ile Gly Gln Gly Ser Tyr Ala Ala Gly Asn Ala Phe  
 1300 1305 1310  
 Leu Asp Ala Leu Ala Asp Leu Arg Arg Thr Gln Gly Leu Ala Ala Leu  
 1315 1320 1325  
 Ser Ile Ala Trp Gly Leu Trp Ala Glu Gly Gly Met Gly Ser Gln Ala  
 1330 1335 1340  
 Gln Arg Arg Glu His Glu Ala Ser Gly Ile Trp Ala Met Pro Thr Ser  
 1345 1350 1355 1360  
 Arg Ala Leu Ala Ala Met Glu Trp Leu Leu Gly Thr Arg Ala Thr Gln  
 1365 1370 1375  
 Arg Val Val Ile Gln Met Asp Trp Ala His Ala Gly Ala Ala Pro Arg  
 1380 1385 1390  
 Asp Ala Ser Arg Gly Arg Phe Trp Asp Arg Leu Val Thr Ala Thr Lys  
 1395 1400 1405  
 Glu Ala Ser Ser Ser Ala Val Pro Ala Val Glu Arg Trp Arg Asn Ala  
 1410 1415 1420  
 Ser Val Val Glu Thr Arg Ser Ala Leu Tyr Glu Leu Val Arg Gly Val  
 1425 1430 1435 1440  
 Val Ala Gly Val Met Gly Phe Thr Asp Gln Gly Thr Leu Asp Val Arg  
 1445 1450 1455  
 Arg Gly Phe Ala Glu Gln Gly Leu Asp Ser Leu Met Ala Val Glu Ile  
 1460 1465 1470  
 Arg Lys Arg Leu Gln Gly Glu Leu Gly Met Pro Leu Ser Ala Thr Leu  
 1475 1480 1485  
 Ala Phe Asp His Pro Thr Val Glu Arg Leu Val Glu Tyr Leu Leu Ser  
 1490 1495 1500  
 Gln Ala Leu Glu Leu Gln Asp Arg Thr Asp Val Arg Ser Val Arg Leu  
 1505 1510 1515 1520  
 Pro Ala Thr Glu Asp Pro Ile Ala Ile Val Gly Ala Ala Cys Arg Phe  
 1525 1530 1535  
 Pro Gly Gly Val Glu Asp Leu Glu Ser Tyr Trp Gln Leu Leu Thr Glu  
 1540 1545 1550  
 Gly Val Val Val Ser Thr Glu Val Pro Ala Asp Arg Trp Asn Gly Ala  
 1555 1560 1565  
 Asp Gly Arg Val Pro Gly Ser Gly Glu Ala Gln Arg Gln Thr Tyr Val  
 1570 1575 1580

Pro Arg Gly Gly Phe Leu Arg Glu Val Glu Thr Phe Asp Ala Ala Phe  
 1585 1590 1595 1600  
 Phe His Ile Ser Pro Arg Glu Ala Met Ser Leu Asp Pro Gln Gln Arg  
 1605 1610 1615  
 Leu Leu Leu Glu Val Ser Trp Glu Ala Ile Glu Arg Ala Gly Gln Asp  
 1620 1625 1630  
 Pro Ser Ala Leu Arg Glu Ser Pro Thr Gly Val Phe Val Gly Ala Gly  
 1635 1640 1645  
 Pro Asn Glu Tyr Ala Glu Arg Val Gln Glu Leu Ala Asp Glu Ala Ala  
 1650 1655 1660  
 Gly Leu Tyr Ser Gly Thr Gly Asn Met Leu Ser Val Ala Ala Gly Arg  
 1665 1670 1675 1680  
 Leu Ser Phe Phe Leu Gly Leu His Gly Pro Thr Leu Ala Val Asp Thr  
 1685 1690 1695  
 Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Gly Cys Gln Ser Leu  
 1700 1705 1710  
 Arg Arg Gly Glu Cys Asp Gln Ala Leu Val Gly Gly Val Asn Met Leu  
 1715 1720 1725  
 Leu Ser Pro Lys Thr Phe Ala Leu Leu Ser Arg Met His Ala Leu Ser  
 1730 1735 1740  
 Pro Gly Gly Arg Cys Lys Thr Phe Ser Ala Asp Ala Asp Gly Tyr Ala  
 1745 1750 1755 1760  
 Arg Ala Glu Gly Cys Ala Val Val Val Leu Lys Arg Leu Ser Asp Ala  
 1765 1770 1775  
 Gln Arg Asp Arg Asp Pro Ile Leu Ala Val Ile Arg Gly Thr Ala Ile  
 1780 1785 1790  
 Asn His Asp Gly Pro Ser Ser Gly Leu Thr Val Pro Ser Gly Pro Ala  
 1795 1800 1805  
 Gln Glu Ala Leu Leu Arg Gln Ala Leu Ala His Ala Gly Val Val Pro  
 1810 1815 1820  
 Ala Asp Val Asp Phe Val Glu Cys His Gly Thr Gly Thr Ala Leu Gly  
 1825 1830 1835 1840  
 Asp Pro Ile Glu Val Arg Ala Leu Ser Asp Val Tyr Gly Gln Ala Arg  
 1845 1850 1855  
 Pro Ala Asp Arg Pro Leu Ile Leu Gly Ala Ala Lys Ala Asn Leu Gly  
 1860 1865 1870  
 His Met Glu Pro Ala Ala Gly Leu Ala Gly Leu Leu Lys Ala Val Leu  
 1875 1880 1885  
 Ala Leu Gly Gln Glu Gln Ile Pro Ala Gln Pro Glu Leu Gly Glu Leu  
 1890 1895 1900  
 Asn Pro Leu Leu Pro Trp Glu Ala Leu Pro Val Ala Val Ala Arg Ala  
 1905 1910 1915 1920  
 Ala Val Pro Trp Pro Arg Thr Asp Arg Pro Arg Phe Ala Gly Val Ser  
 1925 1930 1935

Ser Phe Gly Met Ser Gly Thr Asn Ala His Val Val Leu Glu Glu Ala  
 1940 1945 1950  
 Pro Ala Val Glu Leu Trp Pro Ala Ala Pro Glu Arg Ser Ala Glu Leu  
 1955 1960 1965  
 Leu Val Leu Ser Gly Lys Ser Glu Gly Ala Leu Asp Ala Gln Ala Ala  
 1970 1975 1980  
 Arg Leu Arg Glu His Leu Asp Met His Pro Glu Leu Gly Leu Gly Asp  
 1985 1990 1995 2000  
 Val Ala Phe Ser Leu Ala Thr Thr Arg Ser Ala Met Asn His Arg Leu  
 2005 2010 2015  
 Ala Val Ala Val Thr Ser Arg Glu Gly Leu Leu Ala Ala Leu Ser Ala  
 2020 2025 2030  
 Val Ala Gln Gly Gln Thr Pro Pro Gly Ala Ala Arg Cys Ile Ala Ser  
 2035 2040 2045  
 Ser Ser Arg Gly Lys Leu Ala Phe Leu Phe Thr Gly Gln Gly Ala Gln  
 2050 2055 2060  
 Thr Pro Gly Met Gly Arg Gly Leu Cys Ala Ala Trp Pro Ala Phe Arg  
 2065 2070 2075 2080  
 Glu Ala Phe Asp Arg Cys Val Ala Leu Phe Asp Arg Glu Leu Asp Arg  
 2085 2090 2095  
 Pro Leu Arg Glu Val Met Trp Ala Glu Pro Gly Ser Ala Glu Ser Leu  
 2100 2105 2110  
 Leu Leu Asp Gln Thr Ala Phe Thr Gln Pro Ala Leu Phe Thr Val Glu  
 2115 2120 2125  
 Tyr Ala Leu Thr Ala Leu Trp Arg Ser Trp Gly Val Glu Pro Glu Leu  
 2130 2135 2140  
 Val Ala Gly His Ser Ala Gly Glu Leu Val Ala Ala Cys Val Ala Gly  
 2145 2150 2155 2160  
 Val Phe Ser Leu Glu Asp Gly Val Arg Leu Val Ala Ala Arg Gly Arg  
 2165 2170 2175  
 Leu Met Gln Gly Leu Ser Ala Gly Gly Ala Met Val Ser Leu Gly Ala  
 2180 2185 2190  
 Pro Glu Ala Glu Val Ala Ala Ala Val Ala Pro His Ala Ala Ser Val  
 2195 2200 2205  
 Ser Ile Ala Ala Val Asn Gly Pro Glu Gln Val Val Ile Ala Gly Val  
 2210 2215 2220  
 Glu Gln Ala Val Gln Ala Ile Ala Ala Gly Phe Ala Ala Arg Gly Ala  
 2225 2230 2235 2240  
 Arg Thr Lys Arg Leu His Val Ser His Ala Ser His Ser Pro Leu Met  
 2245 2250 2255  
 Glu Pro Met Leu Glu Glu Phe Gly Arg Val Ala Ala Ser Val Thr Tyr  
 2260 2265 2270  
 Arg Arg Pro Ser Val Ser Leu Val Ser Asn Leu Ser Gly Lys Val Val  
 2275 2280 2285

Ala Asp Glu Leu Ser Ala Pro Gly Tyr Trp Val Arg His Val Arg Glu  
 2290 2295 2300  
 Ala Val Arg Phe Ala Asp Gly Val Lys Ala Leu His Glu Ala Gly Ala  
 2305 2310 2315 2320  
 Gly Thr Phe Val Glu Val Gly Pro Lys Pro Thr Leu Leu Gly Leu Leu  
 2325 2330 2335  
 Pro Ala Cys Leu Pro Glu Ala Glu Pro Thr Leu Leu Ala Ser Leu Arg  
 2340 2345 2350  
 Ala Gly Arg Glu Glu Ala Ala Gly Val Leu Glu Ala Leu Gly Arg Leu  
 2355 2360 2365  
 Trp Ala Ala Gly Gly Ser Val Ser Trp Pro Gly Val Phe Pro Thr Ala  
 2370 2375 2380  
 Gly Arg Arg Val Pro Leu Pro Thr Tyr Pro Trp Gln Arg Gln Arg Tyr  
 2385 2390 2395 2400  
 Trp Pro Asp Ile Glu Pro Asp Ser Arg Arg His Ala Ala Ala Asp Pro  
 2405 2410 2415  
 Thr Gln Gly Trp Phe Tyr Arg Val Asp Trp Pro Glu Ile Pro Arg Ser  
 2420 2425 2430  
 Leu Gln Lys Ser Glu Glu Ala Ser Arg Gly Ser Trp Leu Val Leu Ala  
 2435 2440 2445  
 Asp Lys Gly Gly Val Gly Glu Ala Val Ala Ala Ala Leu Ser Thr Arg  
 2450 2455 2460  
 Gly Leu Pro Cys Val Val Leu His Ala Pro Ala Glu Thr Ser Ala Thr  
 2465 2470 2475 2480  
 Ala Glu Leu Val Thr Glu Ala Ala Gly Gly Arg Ser Asp Trp Gln Val  
 2485 2490 2495  
 Val Leu Tyr Leu Trp Gly Leu Asp Ala Val Val Gly Ala Glu Ala Ser  
 2500 2505 2510  
 Ile Asp Glu Ile Gly Asp Ala Thr Arg Arg Ala Thr Ala Pro Val Leu  
 2515 2520 2525  
 Gly Leu Ala Arg Phe Leu Ser Thr Val Ser Cys Ser Pro Arg Leu Trp  
 2530 2535 2540  
 Val Val Thr Arg Gly Ala Cys Ile Val Gly Asp Glu Pro Ala Ile Ala  
 2545 2550 2555 2560  
 Pro Cys Gln Ala Ala Leu Trp Gly Met Gly Arg Val Ala Ala Leu Glu  
 2565 2570 2575  
 His Pro Gly Ala Trp Gly Gly Leu Val Asp Leu Asp Pro Arg Ala Ser  
 2580 2585 2590  
 Pro Pro Gln Ala Ser Pro Ile Asp Gly Glu Met Leu Val Thr Glu Leu  
 2595 2600 2605  
 Leu Ser Gln Glu Thr Glu Asp Gln Leu Ala Phe Arg His Gly Arg Arg  
 2610 2615 2620  
 His Ala Ala Arg Leu Val Ala Ala Pro Pro Gln Gly Gln Ala Ala Pro  
 2625 2630 2635 2640



Val Ser Leu Ser Ala Glu Ala Ser Tyr Leu Val Thr Gly Gly Leu Gly  
 2645 2650 2655  
 Gly Leu Gly Leu Ile Val Ala Gln Trp Leu Val Glu Leu Gly Ala Arg  
 2660 2665 2670  
 His Leu Val Leu Thr Ser Arg Arg Gly Leu Pro Asp Arg Gln Ala Trp  
 2675 2680 2685  
 Cys Glu Gln Gln Pro Pro Glu Ile Arg Ala Arg Ile Ala Ala Val Glu  
 2690 2695 2700  
 Ala Leu Glu Ala Arg Gly Ala Arg Val Thr Val Ala Ala Val Asp Val  
 2705 2710 2715 2720  
 Ala Asp Val Glu Pro Met Thr Ala Leu Val Ser Ser Val Glu Pro Pro  
 2725 2730 2735  
 Leu Arg Gly Val Val His Ala Ala Gly Val Ser Val Met Arg Pro Leu  
 2740 2745 2750  
 Ala Glu Thr Asp Glu Thr Leu Leu Glu Ser Val Leu Arg Pro Lys Val  
 2755 2760 2765  
 Ala Gly Ser Trp Leu Leu His Arg Leu Leu His Gly Arg Pro Leu Asp  
 2770 2775 2780  
 Leu Phe Val Leu Phe Ser Ser Gly Ala Ala Val Trp Gly Ser His Ser  
 2785 2790 2795 2800  
 Gln Gly Ala Tyr Ala Ala Ala Asn Ala Phe Leu Asp Gly Leu Ala His  
 2805 2810 2815  
 Leu Arg Arg Ser Gln Ser Leu Pro Ala Leu Ser Val Ala Trp Gly Leu  
 2820 2825 2830  
 Trp Ala Glu Gly Gly Met Ala Asp Ala Glu Ala His Ala Arg Leu Ser  
 2835 2840 2845  
 Asp Ile Gly Val Leu Pro Met Ser Thr Ser Ala Ala Leu Ser Ala Leu  
 2850 2855 2860  
 Gln Arg Leu Val Glu Thr Gly Ala Ala Gln Arg Thr Val Thr Arg Met  
 2865 2870 2875 2880  
 Asp Trp Ala Arg Phe Ala Pro Val Tyr Thr Ala Arg Gly Arg Arg Asn  
 2885 2890 2895  
 Leu Leu Ser Ala Leu Val Ala Gly Arg Asp Ile Ile Ala Pro Ser Pro  
 2900 2905 2910  
 Pro Ala Ala Ala Thr Arg Asn Trp Arg Gly Leu Ser Val Ala Glu Ala  
 2915 2920 2925  
 Arg Val Ala Leu His Glu Ile Val His Gly Ala Val Ala Arg Val Leu  
 2930 2935 2940  
 Gly Phe Leu Asp Pro Ser Ala Leu Asp Pro Gly Met Gly Phe Asn Glu  
 2945 2950 2955 2960  
 Gln Gly Leu Asp Ser Leu Met Ala Val Glu Ile Arg Asn Leu Leu Gln  
 2965 2970 2975  
 Ala Glu Leu Asp Val Arg Leu Ser Thr Thr Leu Ala Phe Asp His Pro  
 2980 2985 2990

Thr Val Gln Arg Leu Val Glu His Leu Leu Val Asp Val Leu Lys Leu  
 2995 3000 3005  
 Glu Asp Arg Ser Asp Thr Gln His Val Arg Ser Leu Ala Ser Asp Glu  
 3010 3015 3020  
 Pro Ile Ala Ile Val Gly Ala Ala Cys Arg Phe Pro Gly Gly Val Glu  
 3025 3030 3035 3040  
 Asp Leu Glu Ser Tyr Trp Gln Leu Leu Ala Glu Gly Val Val Val Ser  
 3045 3050 3055  
 Ala Glu Val Pro Ala Asp Arg Trp Asp Ala Ala Asp Trp Tyr Asp Pro  
 3060 3065 3070  
 Asp Pro Glu Ile Pro Gly Arg Thr Tyr Val Thr Lys Gly Ala Phe Leu  
 3075 3080 3085  
 Arg Asp Leu Gln Arg Leu Asp Ala Thr Phe Phe Arg Ile Ser Pro Arg  
 3090 3095 3100  
 Glu Ala Met Ser Leu Asp Pro Gln Gln Arg Leu Leu Leu Glu Val Ser  
 3105 3110 3115 3120  
 Trp Glu Ala Leu Glu Ser Ala Gly Ile Ala Pro Asp Thr Leu Arg Asp  
 3125 3130 3135  
 Ser Pro Thr Gly Val Phe Val Gly Ala Gly Pro Asn Glu Tyr Tyr Thr  
 3140 3145 3150  
 Gln Arg Leu Arg Gly Phe Thr Asp Gly Ala Ala Gly Leu Tyr Gly Gly  
 3155 3160 3165  
 Thr Gly Asn Met Leu Ser Val Thr Ala Gly Arg Leu Ser Phe Phe Leu  
 3170 3175 3180  
 Gly Leu His Gly Pro Thr Leu Ala Met Asp Thr Ala Cys Ser Ser Ser  
 3185 3190 3195 3200  
 Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu Cys  
 3205 3210 3215  
 Asp Gln Ala Leu Val Gly Gly Val Asn Val Leu Leu Ala Pro Glu Thr  
 3220 3225 3230  
 Phe Val Leu Leu Ser Arg Met Arg Ala Leu Ser Pro Asp Gly Arg Cys  
 3235 3240 3245  
 Lys Thr Phe Ser Ala Asp Ala Asp Gly Tyr Ala Arg Gly Glu Gly Cys  
 3250 3255 3260  
 Ala Val Val Val Leu Lys Arg Leu Arg Asp Ala Gln Arg Ala Gly Asp  
 3265 3270 3275 3280  
 Ser Ile Leu Ala Leu Ile Arg Gly Ser Ala Val Asn His Asp Gly Pro  
 3285 3290 3295  
 Ser Ser Gly Leu Thr Val Pro Asn Gly Pro Ala Gln Gln Ala Leu Leu  
 3300 3305 3310  
 Arg Gln Ala Leu Ser Gln Ala Gly Val Ser Pro Val Asp Val Asp Phe  
 3315 3320 3325  
 Val Glu Cys His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu Val  
 3330 3335 3340

Gln Ala Leu Ser Glu Val Tyr Gly Pro Gly Arg Ser Gly Asp Arg Pro  
 3345 3350 3355 3360  
 Leu Val Leu Gly Ala Ala Lys Ala Asn Val Ala His Leu Glu Ala Ala  
 3365 3370 3375  
 Ser Gly Leu Ala Ser Leu Leu Lys Ala Val Leu Ala Leu Arg His Glu  
 3380 3385 3390  
 Gln Ile Pro Ala Gln Pro Glu Leu Gly Glu Leu Asn Pro His Leu Pro  
 3395 3400 3405  
 Trp Asn Thr Leu Pro Val Ala Val Pro Arg Lys Ala Val Pro Trp Gly  
 3410 3415 3420  
 Arg Gly Ala Arg Pro Arg Arg Ala Gly Val Ser Ala Phe Gly Leu Ser  
 3425 3430 3435 3440  
 Gly Thr Asn Val His Val Val Leu Glu Glu Ala Pro Glu Val Glu Pro  
 3445 3450 3455  
 Ala Pro Ala Ala Pro Ala Arg Pro Val Glu Leu Val Val Leu Ser Ala  
 3460 3465 3470  
 Lys Ser Ala Ala Ala Leu Asp Ala Ala Ala Arg Leu Ser Ala His  
 3475 3480 3485  
 Leu Ser Ala His Pro Glu Leu Ser Leu Gly Asp Val Ala Phe Ser Leu  
 3490 3495 3500  
 Ala Thr Thr Arg Ser Pro Met Glu His Arg Leu Ala Ile Ala Thr Thr  
 3505 3510 3515 3520  
 Ser Arg Glu Ala Leu Arg Gly Ala Leu Asp Ala Ala Ala Gln Gln Lys  
 3525 3530 3535  
 Thr Pro Gln Gly Ala Val Arg Gly Lys Ala Val Ser Ser Arg Gly Lys  
 3540 3545 3550  
 Leu Ala Phe Leu Phe Thr Gly Gln Gly Ala Gln Met Pro Gly Met Gly  
 3555 3560 3565  
 Arg Gly Leu Tyr Glu Thr Trp Pro Ala Phe Arg Glu Ala Phe Asp Arg  
 3570 3575 3580  
 Cys Val Ala Leu Phe Asp Arg Glu Ile Asp Gln Pro Leu Arg Glu Val  
 3585 3590 3595 3600  
 Met Trp Ala Ala Pro Gly Leu Ala Gln Ala Ala Arg Leu Asp Gln Thr  
 3605 3610 3615  
 Ala Tyr Ala Gln Pro Ala Leu Phe Ala Leu Glu Tyr Ala Leu Ala Ala  
 3620 3625 3630  
 Leu Trp Arg Ser Trp Gly Val Glu Pro His Val Leu Leu Gly His Ser  
 3635 3640 3645  
 Ile Gly Glu Leu Val Ala Ala Cys Val Ala Gly Val Phe Ser Leu Glu  
 3650 3655 3660  
 Asp Ala Val Arg Leu Val Ala Ala Arg Gly Arg Leu Met Gln Ala Leu  
 3665 3670 3675 3680  
 Pro Ala Gly Gly Ala Met Val Ala Ile Ala Ala Ser Glu Ala Glu Val  
 3685 3690 3695

Ala Ala Ser Val Ala Pro His Ala Ala Thr Val Ser Ile Ala Ala Val  
 3700 3705 3710  
 Asn Gly Pro Asp Ala Val Val Ile Ala Gly Ala Glu Val Gln Val Leu  
 3715 3720 3725  
 Ala Leu Gly Ala Thr Phe Ala Ala Arg Gly Ile Arg Thr Lys Arg Leu  
 3730 3735 3740  
 Ala Val Ser His Ala Phe His Ser Pro Leu Met Asp Pro Met Leu Glu  
 3745 3750 3755 3760  
 Asp Phe Gln Arg Val Ala Ala Thr Ile Ala Tyr Arg Ala Pro Asp Arg  
 3765 3770 3775  
 Pro Val Val Ser Asn Val Thr Gly His Val Ala Gly Pro Glu Ile Ala  
 3780 3785 3790  
 Thr Pro Glu Tyr Trp Val Arg His Val Arg Ser Ala Val Arg Phe Gly  
 3795 3800 3805  
 Asp Gly Ala Lys Ala Leu His Ala Ala Gly Ala Ala Thr Phe Val Glu  
 3810 3815 3820  
 Val Gly Pro Lys Pro Val Leu Leu Gly Leu Leu Pro Ala Cys Leu Gly  
 3825 3830 3835 3840  
 Glu Ala Asp Ala Val Leu Val Pro Ser Leu Arg Ala Asp Arg Ser Glu  
 3845 3850 3855  
 Cys Glu Val Val Leu Ala Ala Leu Gly Ala Trp Tyr Ala Trp Gly Gly  
 3860 3865 3870  
 Ala Leu Asp Trp Lys Gly Val Phe Pro Asp Gly Ala Arg Arg Val Ala  
 3875 3880 3885  
 Leu Pro Met Tyr Pro Trp Gln Arg Glu Arg His Trp Met Asp Leu Thr  
 3890 3895 3900  
 Pro Arg Ser Ala Ala Pro Ala Gly Ile Ala Gly Arg Trp Pro Leu Ala  
 3905 3910 3915 3920  
 Gly Val Gly Leu Cys Met Pro Gly Ala Val Leu His His Val Leu Ser  
 3925 3930 3935  
 Ile Gly Pro Arg His Gln Pro Phe Leu Gly Asp His Leu Val Phe Gly  
 3940 3945 3950  
 Lys Val Val Val Pro Gly Ala Phe His Val Ala Val Ile Leu Ser Ile  
 3955 3960 3965  
 Ala Ala Glu Arg Trp Pro Glu Arg Ala Ile Glu Leu Thr Gly Val Glu  
 3970 3975 3980  
 Phe Leu Lys Ala Ile Ala Met Glu Pro Asp Gln Glu Val Glu Leu His  
 3985 3990 3995 4000  
 Ala Val Leu Thr Pro Glu Ala Ala Gly Asp Gly Tyr Leu Phe Glu Leu  
 4005 4010 4015  
 Ala Thr Leu Ala Ala Pro Glu Thr Glu Arg Arg Trp Thr Thr His Ala  
 4020 4025 4030  
 Arg Gly Arg Val Gln Pro Thr Asp Gly Ala Pro Gly Ala Leu Pro Arg  
 4035 4040 4045

Leu Glu Val Leu Glu Asp Arg Ala Ile Gln Pro Leu Asp Phe Ala Gly  
 4050 4055 4060  
 Phe Leu Asp Arg Leu Ser Ala Val Arg Ile Gly Trp Gly Pro Leu Trp  
 4065 4070 4075 4080  
 Arg Trp Leu Gln Asp Gly Arg Val Gly Asp Glu Ala Ser Leu Ala Thr  
 4085 4090 4095  
 Leu Val Pro Thr Tyr Pro Asn Ala His Asp Val Ala Pro Leu His Pro  
 4100 4105 4110  
 Ile Leu Leu Asp Asn Gly Phe Ala Val Ser Leu Leu Ser Thr Arg Ser  
 4115 4120 4125  
 Glu Pro Glu Asp Asp Gly Thr Pro Pro Leu Pro Phe Ala Val Glu Arg  
 4130 4135 4140  
 Val Arg Trp Trp Arg Ala Pro Val Gly Arg Val Arg Cys Gly Gly Val  
 4145 4150 4155 4160  
 Pro Arg Ser Gln Ala Phe Gly Val Ser Ser Phe Val Leu Val Asp Glu  
 4165 4170 4175  
 Thr Gly Glu Val Val Ala Glu Val Glu Gly Phe Val Cys Arg Arg Ala  
 4180 4185 4190  
 Pro Arg Glu Val Phe Leu Arg Gln Glu Ser Gly Ala Ser Thr Ala Ala  
 4195 4200 4205  
 Leu Tyr Arg Leu Asp Trp Pro Glu Ala Pro Leu Pro Asp Ala Pro Ala  
 4210 4215 4220  
 Glu Arg Ile Glu Glu Ser Trp Val Val Val Ala Ala Pro Gly Ser Glu  
 4225 4230 4235 4240  
 Met Ala Ala Ala Leu Ala Thr Arg Leu Asn Arg Cys Val Leu Ala Glu  
 4245 4250 4255  
 Pro Lys Gly Leu Glu Ala Ala Leu Ala Gly Val Ser Pro Ala Gly Val  
 4260 4265 4270  
 Ile Cys Leu Trp Glu Ala Gly Ala His Glu Glu Ala Pro Ala Ala Ala  
 4275 4280 4285  
 Gln Arg Val Ala Thr Glu Gly Leu Ser Val Val Gln Ala Leu Arg Asp  
 4290 4295 4300  
 Arg Ala Val Arg Leu Trp Trp Val Thr Met Gly Ala Val Ala Val Glu  
 4305 4310 4315 4320  
 Ala Gly Glu Arg Val Gln Val Ala Thr Ala Pro Val Trp Gly Leu Gly  
 4325 4330 4335  
 Arg Thr Val Met Gln Glu Arg Pro Glu Leu Ser Cys Thr Leu Val Asp  
 4340 4345 4350  
 Leu Glu Pro Glu Ala Asp Ala Ala Arg Ser Ala Asp Val Leu Leu Arg  
 4355 4360 4365  
 Glu Leu Gly Arg Ala Asp Asp Glu Thr Gln Val Ala Phe Arg Ser Gly  
 4370 4375 4380  
 Lys Arg Arg Val Ala Arg Leu Val Lys Ala Thr Thr Pro Glu Gly Leu  
 4385 4390 4395 4400

- 45 -

Leu Val Pro Asp Ala Glu Ser Tyr Arg Leu Glu Ala Gly Gln Lys Gly  
 4405 4410 4415  
 Thr Leu Asp Gln Leu Arg Leu Ala Pro Ala Gln Arg Arg Ala Pro Gly  
 4420 4425 4430  
 Pro Gly Glu Val Glu Ile Lys Val Thr Ala Ser Gly Leu Asn Phe Arg  
 4435 4440 4445  
 Thr Val Leu Ala Val Leu Gly Met Tyr Pro Gly Asp Ala Gly Pro Met  
 4450 4455 4460  
 Gly Gly Asp Cys Ala Gly Val Ala Thr Ala Val Gly Gln Gly Val Arg  
 4465 4470 4475 4480  
 His Val Ala Val Gly Asp Ala Val Met Thr Leu Gly Thr Leu His Arg  
 4485 4490 4495  
 Phe Val Thr Val Asp Ala Arg Leu Val Val Arg Gln Pro Ala Gly Leu  
 4500 4505 4510  
 Thr Pro Ala Gln Ala Ala Thr Val Pro Val Ala Phe Leu Thr Ala Trp  
 4515 4520 4525  
 Leu Ala Leu His Asp Leu Gly Asn Leu Arg Arg Gly Glu Arg Val Leu  
 4530 4535 4540  
 Ile His Ala Ala Ala Gly Gly Val Gly Met Ala Ala Val Gln Ile Ala  
 4545 4550 4555 4560  
 Arg Trp Ile Gly Ala Glu Val Phe Ala Thr Ala Ser Pro Ser Lys Trp  
 4565 4570 4575  
 Ala Ala Val Gln Ala Met Gly Val Pro Arg Thr His Ile Ala Ser Ser  
 4580 4585 4590  
 Arg Thr Leu Glu Phe Ala Glu Thr Phe Arg Gln Val Thr Gly Gly Arg  
 4595 4600 4605  
 Gly Val Asp Val Val Leu Asn Ala Leu Ala Gly Glu Phe Val Asp Ala  
 4610 4615 4620  
 Ser Leu Ser Leu Leu Ser Thr Gly Gly Arg Phe Leu Glu Met Gly Lys  
 4625 4630 4635 4640  
 Thr Asp Ile Arg Asp Arg Ala Ala Val Ala Ala Ala His Pro Gly Val  
 4645 4650 4655  
 Arg Tyr Arg Val Phe Asp Ile Leu Glu Leu Ala Pro Asp Arg Thr Arg  
 4660 4665 4670  
 Glu Ile Leu Glu Arg Val Val Glu Gly Phe Ala Ala Gly His Leu Arg  
 4675 4680 4685  
 Ala Leu Pro Val His Ala Phe Ala Ile Thr Lys Ala Glu Ala Ala Phe  
 4690 4695 4700  
 Arg Phe Met Ala Gln Ala Arg His Gln Gly Lys Val Val Leu Leu Pro  
 4705 4710 4715 4720  
 Ala Pro Ser Ala Ala Pro Leu Ala Pro Thr Gly Thr Val Leu Leu Thr  
 4725 4730 4735  
 Gly Gly Leu Gly Ala Leu Gly Leu His Val Ala Arg Trp Leu Ala Gln  
 4740 4745 4750

Gln Gly Val Pro His Met Val Leu Thr Gly Arg Arg Gly Leu Asp Thr  
 4755 4760 4765  
 Pro Gly Ala Ala Lys Ala Val Ala Glu Ile Glu Ala Leu Gly Ala Arg  
 4770 4775 4780  
 Val Thr Ile Ala Ala Ser Asp Val Ala Asp Arg Asn Ala Leu Glu Ala  
 4785 4790 4795 4800  
 Val Leu Gln Ala Ile Pro Ala Glu Trp Pro Leu Gln Gly Val Ile His  
 4805 4810 4815  
 Ala Ala Gly Ala Leu Asp Asp Gly Val Leu Asp Glu Gln Thr Thr Asp  
 4820 4825 4830  
 Arg Phe Ser Arg Val Leu Ala Pro Lys Val Thr Gly Ala Trp Asn Leu  
 4835 4840 4845  
 His Glu Leu Thr Ala Gly Asn Asp Leu Ala Phe Phe Val Leu Phe Ser  
 4850 4855 4860  
 Ser Met Ser Gly Leu Leu Gly Ser Ala Gly Gln Ser Asn Tyr Ala Ala  
 4865 4870 4875 4880  
 Ala Asn Thr Phe Leu Asp Ala Leu Ala Ala His Arg Arg Ala Glu Gly  
 4885 4890 4895  
 Leu Ala Ala Gln Ser Leu Ala Trp Gly Pro Trp Ser Asp Gly Gly Met  
 4900 4905 4910  
 Ala Ala Gly Leu Ser Ala Ala Leu Gln Ala Arg Leu Ala Arg His Gly  
 4915 4920 4925  
 Met Gly Ala Leu Ser Pro Ala Gln Gly Thr Ala Leu Leu Gly Gln Ala  
 4930 4935 4940  
 Leu Ala Arg Pro Glu Thr Gln Leu Gly Ala Met Ser Leu Asp Val Arg  
 4945 4950 4955 4960  
 Ala Ala Ser Gln Ala Ser Gly Ala Ala Val Pro Pro Val Trp Arg Ala  
 4965 4970 4975  
 Leu Val Arg Ala Glu Ala Arg His Thr Ala Ala Gly Ala Gln Gly Ala  
 4980 4985 4990  
 Leu Ala Ala Arg Leu Gly Ala Leu Pro Glu Ala Arg Arg Ala Asp Glu  
 4995 5000 5005  
 Val Arg Lys Val Val Gln Ala Glu Ile Ala Arg Val Leu Ser Trp Ser  
 5010 5015 5020  
 Ala Ala Ser Ala Val Pro Val Asp Arg Pro Leu Ser Asp Leu Gly Leu  
 5025 5030 5035 5040  
 Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Val Leu Gly Gln Arg Val  
 5045 5050 5055  
 Gly Ala Thr Leu Pro Ala Thr Leu Ala Phe Asp His Pro Thr Val Asp  
 5060 5065 5070  
 Ala Leu Thr Arg Trp Leu Leu Asp Lys Val Leu Ala Val Ala Glu Pro  
 5075 5080 5085  
 Ser Val Ser Ser Ala Lys Ser Ser Pro Gln Val Ala Leu Asp Glu Pro  
 5090 5095 5100

Ile Ala Ile Ile Gly Ile Gly Cys Arg Phe Pro Gly Gly Val Ala Asp  
 5105 5110 5115 5120  
 Pro Glu Ser Phe Trp Arg Leu Leu Glu Glu Gly Ser Asp Ala Val Val  
 5125 5130 5135  
 Glu Val Pro His Glu Arg Trp Asp Ile Asp Ala Phe Tyr Asp Pro Asp  
 5140 5145 5150  
 Pro Asp Val Arg Gly Lys Met Thr Thr Arg Phe Gly Gly Phe Leu Ser  
 5155 5160 5165  
 Asp Ile Asp Arg Phe Asp Pro Ala Phe Phe Gly Ile Ser Pro Arg Glu  
 5170 5175 5180  
 Ala Thr Thr Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Ser Trp  
 5185 5190 5195 5200  
 Glu Ala Phe Glu Arg Ala Gly Ile Leu Pro Glu Arg Leu Met Gly Ser  
 5205 5210 5215  
 Asp Thr Gly Val Phe Val Gly Leu Phe Tyr Gln Glu Tyr Ala Ala Leu  
 5220 5225 5230  
 Ala Gly Gly Ile Glu Ala Phe Asp Gly Tyr Leu Gly Thr Gly Thr Thr  
 5235 5240 5245  
 Ala Ser Val Ala Ser Gly Arg Ile Ser Tyr Val Leu Gly Leu Lys Gly  
 5250 5255 5260  
 Pro Ser Leu Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Val  
 5265 5270 5275 5280  
 His Leu Ala Cys Gln Ala Leu Arg Arg Gly Glu Cys Ser Val Ala Leu  
 5285 5290 5295  
 Ala Gly Gly Val Ala Leu Met Leu Thr Pro Ala Thr Phe Val Glu Phe  
 5300 5305 5310  
 Ser Arg Leu Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ser  
 5315 5320 5325  
 Ala Ala Ala Asp Gly Val Gly Trp Ser Glu Gly Cys Ala Met Leu Leu  
 5330 5335 5340  
 Leu Lys Pro Leu Arg Asp Ala Gln Arg Asp Gly Asp Pro Ile Leu Ala  
 5345 5350 5355 5360  
 Val Ile Arg Gly Thr Ala Val Asn Gln Asp Gly Arg Ser Asn Gly Leu  
 5365 5370 5375  
 Thr Ala Pro Asn Gly Ser Ser Gln Gln Glu Val Ile Arg Arg Ala Leu  
 5380 5385 5390  
 Glu Gln Ala Gly Leu Ala Pro Ala Asp Val Ser Tyr Val Glu Cys His  
 5395 5400 5405  
 Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Val Gln Ala Leu Gly  
 5410 5415 5420  
 Ala Val Leu Ala Gln Gly Arg Pro Ser Asp Arg Pro Leu Val Ile Gly  
 5425 5430 5435 5440  
 Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala  
 5445 5450 5455



Gly Val Ile Lys Val Ala Leu Ala Leu Glu Arg Gly Leu Ile Pro Arg  
 5460 5465 5470  
 Ser Leu His Phe Asp Ala Pro Asn Pro His Ile Pro Trp Ser Glu Leu  
 5475 5480 5485  
 Ala Val Gln Val Ala Ala Lys Pro Val Glu Trp Thr Arg Asn Gly Val  
 5490 5495 5500  
 Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala  
 5505 5510 5515 5520  
 His Val Val Leu Glu Ala Pro Ala Ala Ala Phe Ala Pro Ala Ala  
 5525 5530 5535  
 Ala Arg Ser Ala Glu Leu Phe Val Leu Ser Ala Lys Ser Ala Ala Ala  
 5540 5545 5550  
 Leu Asp Ala Gln Ala Ala Arg Leu Ser Ala His Val Val Ala His Pro  
 5555 5560 5565  
 Glu Leu Gly Leu Gly Asp Leu Ala Phe Ser Leu Ala Thr Thr Arg Ser  
 5570 5575 5580  
 Pro Met Thr Tyr Arg Leu Ala Val Ala Ala Thr Ser Arg Glu Ala Leu  
 5585 5590 5595 5600  
 Ser Ala Ala Leu Asp Thr Ala Ala Gln Gly Gln Ala Pro Pro Ala Ala  
 5605 5610 5615  
 Ala Arg Gly His Ala Ser Thr Gly Ser Ala Pro Lys Val Val Phe Val  
 5620 5625 5630  
 Phe Pro Gly Gln Gly Ser Gln Trp Leu Gly Met Gly Gln Lys Leu Leu  
 5635 5640 5645  
 Ser Glu Glu Pro Val Phe Arg Asp Ala Leu Ser Ala Cys Asp Arg Ala  
 5650 5655 5660  
 Ile Gln Ala Glu Ala Gly Trp Ser Leu Leu Ala Glu Leu Ala Ala Asp  
 5665 5670 5675 5680  
 Glu Thr Thr Ser Gln Leu Gly Arg Ile Asp Val Val Gln Pro Ala Leu  
 5685 5690 5695  
 Phe Ala Ile Glu Val Ala Leu Ser Ala Leu Trp Arg Ser Trp Gly Val  
 5700 5705 5710  
 Glu Pro Asp Ala Val Val Gly His Ser Met Gly Glu Val Ala Ala Ala  
 5715 5720 5725  
 His Val Ala Gly Ala Leu Ser Leu Glu Asp Ala Val Ala Ile Ile Cys  
 5730 5735 5740  
 Arg Arg Ser Leu Leu Leu Arg Arg Ile Ser Gly Gln Gly Glu Met Ala  
 5745 5750 5755 5760  
 Val Val Glu Leu Ser Leu Ala Glu Ala Glu Ala Ala Leu Leu Gly Tyr  
 5765 5770 5775  
 Glu Asp Arg Leu Ser Val Ala Val Ser Asn Ser Pro Arg Ser Thr Val  
 5780 5785 5790  
 Leu Ala Gly Glu Pro Ala Ala Leu Ala Glu Val Leu Ala Ile Leu Ala  
 5795 5800 5805

Ala Lys Gly Val Phe Cys Arg Arg Val Lys Val Asp Val Ala Ser His  
 5810 5815 5820  
 Ser Pro Gln Ile Asp Pro Leu Arg Asp Glu Leu Leu Ala Ala Leu Gly  
 5825 5830 5835 5840  
 Glu Leu Glu Pro Arg Gln Ala Thr Val Ser Met Arg Ser Thr Val Thr  
 5845 5850 5855  
 Ser Thr Ile Met Ala Gly Pro Glu Leu Val Ala Ser Tyr Trp Ala Asp  
 5860 5865 5870  
 Asn Val Arg Gln Pro Val Arg Phe Ala Glu Ala Val Gln Ser Leu Met  
 5875 5880 5885  
 Glu Asp Gly His Gly Leu Phe Val Glu Met Ser Pro His Pro Ile Leu  
 5890 5895 5900  
 Thr Thr Ser Val Glu Glu Ile Arg Arg Ala Thr Lys Arg Glu Gly Val  
 5905 5910 5915 5920  
 Ala Val Gly Ser Leu Arg Arg Gly Gln Asp Glu Arg Leu Ser Met Leu  
 5925 5930 5935  
 Glu Ala Leu Gly Ala Leu Trp Val His Gly Gln Ala Val Gly Trp Glu  
 5940 5945 5950  
 Arg Leu Phe Ser Ala Gly Gly Ala Gly Leu Arg Arg Val Pro Leu Pro  
 5955 5960 5965  
 Thr Tyr Pro Trp Gln Arg Glu Arg Tyr Trp Val Asp Ala Pro Thr Gly  
 5970 5975 5980  
 Gly Ala Ala Gly Gly Ser Arg Phe Ala His Ala Gly Ser His Pro Leu  
 5985 5990 5995 6000  
 Leu Gly Glu Met Gln Thr Leu Ser Thr Gln Arg Ser Thr Arg Val Trp  
 6005 6010 6015  
 Glu Thr Thr Leu Asp Leu Lys Arg Leu Pro Trp Leu Gly Asp His Arg  
 6020 6025 6030  
 Val Gln Gly Ala Val Val Phe Pro Gly Ala Ala Tyr Leu Glu Met Ala  
 6035 6040 6045  
 Leu Ser Ser Gly Ala Glu Ala Leu Gly Asp Gly Pro Leu Gln Val Ser  
 6050 6055 6060  
 Asp Val Val Leu Ala Glu Ala Leu Ala Phe Ala Asp Asp Thr Pro Ala  
 6065 6070 6075 6080  
 Ala Val Gln Val Met Ala Thr Glu Glu Arg Pro Gly Arg Leu Gln Phe  
 6085 6090 6095  
 His Val Ala Ser Arg Val Pro Gly His Gly Gly Ala Ala Phe Arg Ser  
 6100 6105 6110  
 His Ala Arg Gly Val Leu Arg Gln Ile Glu Arg Ala Glu Val Pro Ala  
 6115 6120 6125  
 Arg Leu Asp Leu Ala Ala Leu Arg Ala Arg Leu Gln Ala Ser Ala Pro  
 6130 6135 6140  
 Ala Ala Ala Thr Tyr Ala Ala Leu Ala Glu Met Gly Leu Glu Tyr Gly  
 6145 6150 6155 6160

Pro Ala Phe Gln Gly Leu Val Glu Leu Trp Arg Gly Glu Gly Glu Ala  
 6165 6170 6175  
 Leu Gly Arg Val Arg Leu Pro Glu Ala Ala Gly Ser Pro Ala Ala Cys  
 6180 6185 6190  
 Arg Leu His Pro Ala Leu Leu Asp Ala Cys Phe His Val Ser Ser Ala  
 6195 6200 6205  
 Phe Ala Asp Arg Gly Glu Ala Thr Pro Trp Val Pro Val Glu Ile Gly  
 6210 6215 6220  
 Ser Leu Arg Trp Phe Gln Arg Pro Ser Gly Glu Leu Trp Cys His Ala  
 6225 6230 6235 6240  
 Arg Ser Val Ser His Gly Lys Pro Thr Pro Asp Arg Arg Ser Thr Asp  
 6245 6250 6255  
 Phe Trp Val Val Asp Ser Thr Gly Ala Ile Val Ala Glu Ile Ser Gly  
 6260 6265 6270  
 Leu Val Ala Gln Arg Leu Ala Gly Gly Val Arg Arg Arg Glu Glu Asp  
 6275 6280 6285  
 Asp Trp Phe Met Glu Pro Ala Trp Glu Pro Thr Ala Val Pro Gly Ser  
 6290 6295 6300  
 Glu Val Met Ala Gly Arg Trp Leu Leu Ile Gly Ser Gly Gly Gly Leu  
 6305 6310 6315 6320  
 Gly Ala Ala Leu His Ser Ala Leu Thr Glu Ala Gly His Ser Val Val  
 6325 6330 6335  
 His Ala Thr Gly Arg Gly Thr Ser Ala Ala Gly Leu Gln Ala Leu Leu  
 6340 6345 6350  
 Thr Ala Ser Phe Asp Gly Gln Ala Pro Thr Ser Val Val His Leu Gly  
 6355 6360 6365  
 Ser Leu Asp Glu Arg Gly Val Leu Asp Ala Asp Ala Pro Phe Asp Ala  
 6370 6375 6380  
 Asp Ala Leu Glu Glu Ser Leu Val Arg Gly Cys Asp Ser Val Leu Trp  
 6385 6390 6395 6400  
 Thr Val Gln Ala Val Ala Gly Ala Gly Phe Arg Asp Pro Pro Arg Leu  
 6405 6410 6415  
 Trp Leu Val Thr Arg Gly Ala Gln Ala Ile Gly Ala Gly Asp Val Ser  
 6420 6425 6430  
 Val Ala Gln Ala Pro Leu Leu Gly Leu Gly Arg Val Ile Ala Leu Glu  
 6435 6440 6445  
 His Ala Glu Leu Arg Cys Ala Arg Ile Asp Leu Asp Pro Ala Arg Arg  
 6450 6455 6460  
 Asp Gly Glu Val Asp Glu Leu Leu Ala Glu Leu Leu Ala Asp Asp Ala  
 6465 6470 6475 6480  
 Glu Glu Glu Val Ala Phe Arg Gly Gly Glu Arg Arg Val Ala Arg Leu  
 6485 6490 6495  
 Val Arg Arg Leu Pro Glu Thr Asp Cys Arg Glu Lys Ile Glu Pro Ala  
 6500 6505 6510

Glu Gly Arg Pro Phe Arg Leu Glu Ile Asp Gly Ser Gly Val Leu Asp  
 6515 6520 6525  
 Asp Leu Val Leu Arg Ala Thr Glu Arg Arg Pro Pro Gly Pro Gly Glu  
 6530 6535 6540  
 Val Glu Ile Ala Val Glu Ala Ala Gly Leu Asn Phe Leu Asp Val Met  
 6545 6550 6555 6560  
 Arg Ala Met Gly Ile Tyr Pro Gly Pro Gly Asp Gly Pro Val Ala Leu  
 6565 6570 6575  
 Gly Ala Glu Cys Ser Gly Arg Ile Val Ala Met Gly Glu Gly Val Glu  
 6580 6585 6590  
 Ser Leu Arg Ile Gly Gln Asp Val Val Ala Val Ala Pro Phe Ser Phe  
 6595 6600 6605  
 Gly Thr His Val Thr Ile Asp Ala Arg Met Leu Ala Pro Arg Pro Ala  
 6610 6615 6620  
 Ala Leu Thr Ala Ala Gln Ala Ala Ala Leu Pro Val Ala Phe Met Thr  
 6625 6630 6635 6640  
 Ala Trp Tyr Gly Leu Val His Leu Gly Arg Leu Arg Ala Gly Glu Arg  
 6645 6650 6655  
 Val Leu Ile His Ser Ala Thr Gly Gly Thr Gly Leu Ala Ala Val Gln  
 6660 6665 6670  
 Ile Ala Arg His Leu Gly Ala Glu Ile Phe Ala Thr Ala Gly Thr Pro  
 6675 6680 6685  
 Glu Lys Arg Ala Trp Leu Arg Glu Gln Gly Ile Ala His Val Met Asp  
 6690 6695 6700  
 Ser Arg Ser Leu Asp Phe Ala Glu Gln Val Leu Ala Ala Thr Lys Gly  
 6705 6710 6715 6720  
 Glu Gly Val Asp Val Val Leu Asn Ser Leu Ser Gly Ala Ala Ile Asp  
 6725 6730 6735  
 Ala Ser Leu Ser Thr Leu Val Pro Asp Gly Arg Phe Ile Glu Leu Gly  
 6740 6745 6750  
 Lys Thr Asp Ile Tyr Ala Asp Arg Ser Leu Gly Leu Ala His Phe Arg  
 6755 6760 6765  
 Lys Ser Leu Ser Tyr Ser Ala Val Asp Leu Ala Gly Leu Ala Val Arg  
 6770 6775 6780  
 Arg Pro Glu Arg Val Ala Ala Leu Leu Ala Glu Val Val Asp Leu Leu  
 6785 6790 6795 6800  
 Ala Arg Gly Ala Leu Gln Pro Leu Pro Val Glu Ile Phe Pro Leu Ser  
 6805 6810 6815  
 Arg Ala Ala Asp Ala Phe Arg Lys Met Ala Gln Ala Gln His Leu Gly  
 6820 6825 6830  
 Lys Leu Val Leu Ala Leu Glu Asp Pro Asp Val Arg Ile Arg Val Pro  
 6835 6840 6845  
 Gly Glu Ser Gly Val Ala Ile Arg Ala Asp Gly Ala Tyr Leu Val Thr  
 6850 6855 6860

Gly Gly Leu Gly Gly Leu Gly Leu Ser Val Ala Gly Trp Leu Ala Glu  
 6865 6870 6875 6880  
 Gln Gly Ala Gly His Leu Val Leu Val Gly Arg Ser Gly Ala Val Ser  
 6885 6890 6895  
 Ala Glu Gln Gln Thr Ala Val Ala Ala Leu Glu Ala His Gly Ala Arg  
 6900 6905 6910  
 Val Thr Val Ala Arg Ala Asp Val Ala Asp Arg Ala Gln Met Glu Arg  
 6915 6920 6925  
 Ile Leu Arg Glu Val Thr Ala Ser Gly Met Pro Leu Arg Gly Val Val  
 6930 6935 6940  
 His Ala Ala Gly Ile Leu Asp Asp Gly Leu Leu Met Gln Gln Thr Pro  
 6945 6950 6955 6960  
 Ala Arg Phe Arg Ala Val Met Ala Pro Lys Val Arg Gly Ala Leu His  
 6965 6970 6975  
 Leu His Ala Leu Thr Arg Glu Ala Pro Leu Ser Phe Phe Val Leu Tyr  
 6980 6985 6990  
 Ala Ser Gly Ala Gly Leu Leu Gly Ser Pro Gly Gln Gly Asn Tyr Ala  
 6995 7000 7005  
 Ala Ala Asn Thr Phe Leu Asp Ala Leu Ala His His Arg Arg Ala Gln  
 7010 7015 7020  
 Gly Leu Pro Ala Leu Ser Ile Asp Trp Gly Leu Phe Ala Asp Val Gly  
 7025 7030 7035 7040  
 Leu Ala Ala Gly Gln Gln Asn Arg Gly Ala Arg Leu Val Thr Arg Gly  
 7045 7050 7055  
 Thr Arg Ser Leu Thr Pro Asp Glu Gly Leu Trp Ala Leu Glu Arg Leu  
 7060 7065 7070  
 Leu Asp Gly Asp Arg Thr Gln Ala Gly Val Met Pro Phe Asp Val Arg  
 7075 7080 7085  
 Gln Trp Val Glu Phe Tyr Pro Ala Ala Ala Ser Ser Arg Arg Leu Ser  
 7090 7095 7100  
 Arg Leu Met Thr Ala Arg Arg Val Ala Ser Gly Arg Leu Ala Gly Asp  
 7105 7110 7115 7120  
 Arg Asp Leu Leu Glu Arg Leu Ala Thr Ala Glu Ala Gly Ala Arg Ala  
 7125 7130 7135  
 Gly Met Leu Gln Glu Val Val Arg Ala Gln Val Ser Gln Val Leu Arg  
 7140 7145 7150  
 Leu Ser Glu Gly Lys Leu Asp Val Asp Ala Pro Leu Thr Ser Leu Gly  
 7155 7160 7165  
 Met Asp Ser Leu Met Gly Leu Glu Leu Arg Asn Arg Ile Glu Ala Val  
 7170 7175 7180  
 Leu Gly Ile Thr Met Pro Ala Thr Leu Leu Trp Thr Tyr Pro Thr Val  
 7185 7190 7195 7200  
 Ala Ala Leu Ser Ala His Leu Ala Ser His Val Val Ser Thr Gly Asp  
 7205 7210 7215

- 53 -

Gly Glu Ser Ala Arg Pro Pro Asp Thr Gly Ser Val Ala Pro Thr Thr  
7220 7225 7230

His Glu Val Ala Ser Leu Asp Glu Asp Gly Leu Phe Ala Leu Ile Asp  
7235 7240 7245

Glu Ser Leu Ala Arg Ala Gly Lys Arg  
7250 7255

<210> 6  
<211> 3798  
<212> PRT  
<213> Sorangium cellulosum

<400> 6  
Val Thr Asp Arg Glu Gly Gln Leu Leu Glu Arg Leu Arg Glu Val Thr  
1 5 10 15

Leu Ala Leu Arg Lys Thr Leu Asn Glu Arg Asp Thr Leu Glu Leu Glu  
20 25 30

Lys Thr Glu Pro Ile Ala Ile Val Gly Ile Gly Cys Arg Phe Pro Gly  
35 40 45

Gly Ala Gly Thr Pro Glu Ala Phe Trp Glu Leu Leu Asp Asp Gly Arg  
50 55 60

Asp Ala Ile Arg Pro Leu Glu Glu Arg Trp Ala Leu Val Gly Val Asp  
65 70 75 80

Pro Gly Asp Asp Val Pro Arg Trp Ala Gly Leu Leu Thr Glu Ala Ile  
85 90 95

Asp Gly Phe Asp Ala Ala Phe Phe Gly Ile Ala Pro Arg Glu Ala Arg  
100 105 110

Ser Leu Asp Pro Gln His Arg Leu Leu Leu Glu Val Ala Trp Glu Gly  
115 120 125

Phe Glu Asp Ala Gly Ile Pro Pro Arg Ser Leu Val Gly Ser Arg Thr  
130 135 140

Gly Val Phe Val Gly Val Cys Ala Thr Glu Tyr Leu His Ala Ala Val  
145 150 155 160

Ala His Gln Pro Arg Glu Glu Arg Asp Ala Tyr Ser Thr Thr Gly Asn  
165 170 175

Met Leu Ser Ile Ala Ala Gly Arg Leu Ser Tyr Thr Leu Gly Leu Gln  
180 185 190

Gly Pro Cys Leu Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala  
195 200 205

Ile His Leu Ala Cys Arg Ser Leu Arg Ala Arg Glu Ser Asp Leu Ala  
210 215 220

Leu Ala Gly Gly Val Asn Met Leu Leu Ser Pro Asp Thr Met Arg Ala  
225 230 235 240

Leu Ala Arg Thr Gln Ala Leu Ser Pro Asn Gly Arg Cys Gln Thr Phe  
245 250 255

Asp Ala Ser Ala Asn Gly Phe Val Arg Gly Glu Gly Cys Gly Leu Ile

- 54 -

260					265					270					
Val	Leu	Lys	Arg	Leu	Ser	Asp	Ala	Arg	Arg	Asp	Gly	Asp	Arg	Ile	Trp
		275					280					285			
Ala	Leu	Ile	Arg	Gly	Ser	Ala	Ile	Asn	Gln	Asp	Gly	Arg	Ser	Thr	Gly
		290				295					300				
Leu	Thr	Ala	Pro	Asn	Val	Leu	Ala	Gln	Gly	Ala	Leu	Leu	Arg	Glu	Ala
305					310					315					320
Leu	Arg	Asn	Ala	Gly	Val	Glu	Ala	Glu	Ala	Ile	Gly	Tyr	Ile	Glu	Thr
				325					330					335	
His	Gly	Ala	Ala	Thr	Ser	Leu	Gly	Asp	Pro	Ile	Glu	Ile	Glu	Ala	Leu
			340					345					350		
Arg	Ala	Val	Val	Gly	Pro	Ala	Arg	Ala	Asp	Gly	Ala	Arg	Cys	Val	Leu
		355					360					365			
Gly	Ala	Val	Lys	Thr	Asn	Leu	Gly	His	Leu	Glu	Gly	Ala	Ala	Gly	Val
		370				375					380				
Ala	Gly	Leu	Ile	Lys	Ala	Thr	Leu	Ser	Leu	His	His	Glu	Arg	Ile	Pro
385					390					395					400
Arg	Asn	Leu	Asn	Phe	Arg	Thr	Leu	Asn	Pro	Arg	Ile	Arg	Ile	Glu	Gly
				405					410					415	
Thr	Ala	Leu	Ala	Leu	Ala	Thr	Glu	Pro	Val	Pro	Trp	Pro	Arg	Thr	Gly
			420					425					430		
Arg	Thr	Arg	Phe	Ala	Gly	Val	Ser	Ser	Phe	Gly	Met	Ser	Gly	Thr	Asn
		435					440					445			
Ala	His	Val	Val	Leu	Glu	Glu	Ala	Pro	Ala	Val	Glu	Pro	Glu	Ala	Ala
		450				455					460				
Ala	Pro	Glu	Arg	Ala	Ala	Glu	Leu	Phe	Val	Leu	Ser	Ala	Lys	Ser	Ala
465					470					475					480
Ala	Ala	Leu	Asp	Ala	Gln	Ala	Ala	Arg	Leu	Arg	Asp	His	Leu	Glu	Lys
				485					490					495	
His	Val	Glu	Leu	Gly	Leu	Gly	Asp	Val	Ala	Phe	Ser	Leu	Ala	Thr	Thr
			500					505					510		
Arg	Ser	Ala	Met	Glu	His	Arg	Leu	Ala	Val	Ala	Ala	Ser	Ser	Arg	Glu
		515					520					525			
Ala	Leu	Arg	Gly	Ala	Leu	Ser	Ala	Ala	Ala	Gln	Gly	His	Thr	Pro	Pro
		530				535					540				
Gly	Ala	Val	Arg	Gly	Arg	Ala	Ser	Gly	Gly	Ser	Ala	Pro	Lys	Val	Val
545					550					555					560
Phe	Val	Phe	Pro	Gly	Gln	Gly	Ser	Gln	Trp	Val	Gly	Met	Gly	Arg	Lys
				565					570					575	
Leu	Met	Ala	Glu	Glu	Pro	Val	Phe	Arg	Ala	Ala	Leu	Glu	Gly	Cys	Asp
			580					585					590		
Arg	Ala	Ile	Glu	Ala	Glu	Ala	Gly	Trp	Ser						

610					615					620					
Val 625	Leu	Phe	Ala	Met	Glu 630	Val	Ala	Leu	Ser	Ala 635	Leu	Trp	Arg	Ser	Trp 640
Gly	Val	Glu	Pro	Glu 645	Ala	Val	Val	Gly	His 650	Ser	Met	Gly	Glu	Val	Ala 655
Ala	Ala	His	Val 660	Ala	Gly	Ala	Leu	Ser 665	Leu	Glu	Asp	Ala	Val 670	Ala	Ile
Ile	Cys	Arg 675	Arg	Ser	Arg	Leu	Leu 680	Arg	Arg	Ile	Ser	Gly 685	Gln	Gly	Glu
Met	Ala 690	Leu	Val	Glu	Leu	Ser 695	Leu	Glu	Glu	Ala	Glu 700	Ala	Ala	Leu	Arg
Gly 705	His	Glu	Gly	Arg	Leu 710	Ser	Val	Ala	Val	Ser 715	Asn	Ser	Pro	Arg	Ser 720
Thr	Val	Leu	Ala	Gly 725	Glu	Pro	Ala	Ala	Leu 730	Ser	Glu	Val	Leu	Ala 735	Ala
Leu	Thr	Ala	Lys 740	Gly	Val	Phe	Trp	Arg 745	Gln	Val	Lys	Val	Asp 750	Val	Ala
Ser	His	Ser 755	Pro	Gln	Val	Asp	Pro 760	Leu	Arg	Glu	Glu	Leu 765	Ile	Ala	Ala
Leu	Gly 770	Ala	Ile	Arg	Pro	Arg 775	Ala	Ala	Ala	Val	Pro 780	Met	Arg	Ser	Thr
Val 785	Thr	Gly	Gly	Val	Ile 790	Ala	Gly	Pro	Glu	Leu 795	Gly	Ala	Ser	Tyr	Trp 800
Ala	Asp	Asn	Leu	Arg 805	Gln	Pro	Val	Arg	Phe 810	Ala	Ala	Ala	Ala	Gln 815	Ala
Leu	Leu	Glu	Gly 820	Gly	Pro	Ala	Leu	Phe 825	Ile	Glu	Met	Ser	Pro 830	His	Pro
Ile	Leu	Val 835	Pro	Pro	Leu	Asp	Glu 840	Ile	Gln	Thr	Ala	Ala 845	Glu	Gln	Gly
Gly	Ala 850	Ala	Val	Gly	Ser	Leu 855	Arg	Arg	Gly	Gln	Asp 860	Glu	Arg	Ala	Thr
Leu 865	Leu	Glu	Ala	Leu	Gly 870	Thr	Leu	Trp	Ala	Ser 875	Gly	Tyr	Pro	Val	Ser 880
Trp	Ala	Arg	Leu	Phe 885	Pro	Ala	Gly	Gly	Arg 890	Arg	Val	Pro	Leu	Pro 895	Thr
Tyr	Pro	Trp	Gln 900	His	Glu	Arg	Cys	Trp 905	Ile	Glu	Val	Glu	Pro 910	Asp	Ala
Arg	Arg	Leu 915	Ala	Ala	Ala	Asp	Pro 920	Thr	Lys	Asp	Trp	Phe 925	Tyr	Arg	Thr
Asp	Trp 930	Pro	Glu	Val	Pro	Arg 935	Ala	Ala	Pro	Lys	Ser 940	Glu	Thr	Ala	His
Gly 945	Ser	Trp	Leu	Leu	Leu 950	Ala	Asp	Arg	Gly	Gly 955	Val	Gly	Glu	Ala	Val 960
Ala	Ala	Ala	Leu	Ser	Thr	Arg	Gly	Leu	Ser	Cys	Thr	Val	Leu	His	Ala



9

1315	1320	1325
Gly Leu Trp Ala Glu Gly Gly Met Val Asp Ala Lys Ala His Ala Arg 1330	1335	1340
Leu Ser Asp Ile Gly Val Leu Pro Met Ala Thr Gly Pro Ala Leu Ser 1345	1350	1355 1360
Ala Leu Glu Arg Leu Val Asn Thr Ser Ala Val Gln Arg Ser Val Thr 1365	1370	1375
Arg Met Asp Trp Ala Arg Phe Ala Pro Val Tyr Ala Ala Arg Gly Arg 1380	1385	1390
Arg Asn Leu Leu Ser Ala Leu Val Ala Glu Asp Glu Arg Ala Ala Ser 1395	1400	1405
Pro Pro Val Pro Thr Ala Asn Arg Ile Trp Arg Gly Leu Ser Val Ala 1410	1415	1420
Glu Ser Arg Ser Ala Leu Tyr Glu Leu Val Arg Gly Ile Val Ala Arg 1425	1430	1435 1440
Val Leu Gly Phe Ser Asp Pro Gly Ala Leu Asp Val Gly Arg Gly Phe 1445	1450	1455
Ala Glu Gln Gly Leu Asp Ser Leu Met Ala Leu Glu Ile Arg Asn Arg 1460	1465	1470
Leu Gln Arg Glu Leu Gly Glu Arg Leu Ser Ala Thr Leu Ala Phe Asp 1475	1480	1485
His Pro Thr Val Glu Arg Leu Val Ala His Leu Leu Thr Asp Val Leu 1490	1495	1500
Lys Leu Glu Asp Arg Ser Asp Thr Arg His Ile Arg Ser Val Ala Ala 1505	1510	1515 1520
Asp Asp Asp Ile Ala Ile Val Gly Ala Ala Cys Arg Phe Pro Gly Gly 1525	1530	1535
Asp Glu Gly Leu Glu Thr Tyr Trp Arg His Leu Ala Glu Gly Met Val 1540	1545	1550
Val Ser Thr Glu Val Pro Ala Asp Arg Trp Arg Ala Ala Asp Trp Tyr 1555	1560	1565
Asp Pro Asp Pro Glu Val Pro Gly Arg Thr Tyr Val Ala Lys Gly Ala 1570	1575	1580
Phe Leu Arg Asp Val Arg Ser Leu Asp Ala Ala Phe Phe Ala Ile Ser 1585	1590	1595 1600
Pro Arg Glu Ala Met Ser Leu Asp Pro Gln Gln Arg Leu Leu Leu Glu 1605	1610	1615
Val Ser Trp Glu Ala Ile Glu Arg Ala Gly Gln Asp Pro Met Ala Leu 1620	1625	1630
Arg Glu Ser Ala Thr Gly Val Phe Val Gly Met Ile Gly Ser Glu His 1635	1640	1645
Ala Glu Arg Val Gln Gly Leu Asp Asp Asp Ala Ala Leu Leu Tyr Gly 1650	1655	1660
Thr Thr Gly Asn Leu Leu Ser Val Ala Ala Gly Arg Leu Ser Phe Phe		

1665	1670	1675	1680
Leu Gly Leu His Gly Pro Thr Met Thr Val Asp Thr Ala Cys Ser Ser	1685	1690	1695
Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu	1700	1705	1710
Cys Asp Gln Ala Leu Ala Gly Gly Ser Ser Val Leu Leu Ser Pro Arg	1715	1720	1725
Ser Phe Val Ala Ala Ser Arg Met Arg Leu Leu Ser Pro Asp Gly Arg	1730	1735	1740
Cys Lys Thr Phe Ser Ala Ala Ala Asp Gly Phe Ala Arg Ala Glu Gly	1745	1750	1755
Cys Ala Val Val Val Leu Lys Arg Leu Arg Asp Ala Gln Arg Asp Arg	1765	1770	1775
Asp Pro Ile Leu Ala Val Val Arg Ser Thr Ala Ile Asn His Asp Gly	1780	1785	1790
Pro Ser Ser Gly Leu Thr Val Pro Ser Gly Pro Ala Gln Gln Ala Leu	1795	1800	1805
Leu Arg Gln Ala Leu Ala Gln Ala Gly Val Ala Pro Ala Glu Val Asp	1810	1815	1820
Phe Val Glu Cys His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu	1825	1830	1835
Val Gln Ala Leu Gly Ala Val Tyr Gly Arg Gly Arg Pro Ala Glu Arg	1845	1850	1855
Pro Leu Trp Leu Gly Ala Val Lys Ala Asn Leu Gly His Leu Glu Ala	1860	1865	1870
Ala Ala Gly Leu Ala Gly Val Leu Lys Val Leu Leu Ala Leu Glu His	1875	1880	1885
Glu Gln Ile Pro Ala Gln Pro Glu Leu Asp Glu Leu Asn Pro His Ile	1890	1895	1900
Pro Trp Ala Glu Leu Pro Val Ala Val Val Arg Arg Ala Val Pro Trp	1905	1910	1915
Pro Arg Gly Ala Arg Pro Arg Arg Ala Gly Val Ser Ala Phe Gly Leu	1925	1930	1935
Ser Gly Thr Asn Ala His Val Val Leu Glu Glu Ala Pro Ala Val Glu	1940	1945	1950
Pro Val Ala Ala Ala Pro Glu Arg Ala Ala Glu Leu Phe Val Leu Ser	1955	1960	1965
Ala Lys Ser Ala Ala Ala Leu Asp Ala Gln Ala Ala Arg Leu Arg Asp	1970	1975	1980
His Leu Glu Lys His Val Glu Leu Gly Leu Gly Asp Val Ala Phe Ser	1985	1990	1995
Leu Ala Thr Thr Arg Ser Ala Met Glu His Arg Leu Ala Val Ala Ala	2005	2010	2015
Ser Ser Arg Glu Ala Leu Arg Gly Ala Leu Ser Ala Ala Ala Gln Gly			

2020	2025	2030
His Thr Pro Pro Gly Ala Val Arg Gly Arg Ala Ser Gly Gly Ser Ala 2035 2040 2045		
Pro Lys Val Val Phe Val Phe Pro Gly Gln Gly Ser Gln Trp Val Gly 2050 2055 2060		
Met Gly Arg Lys Leu Met Ala Glu Glu Pro Val Phe Arg Ala Ala Leu 2065 2070 2075 2080		
Glu Gly Cys Asp Arg Ala Ile Glu Ala Glu Ala Gly Trp Ser Leu Leu 2085 2090 2095		
Gly Glu Leu Ser Ala Asp Glu Ala Ala Ser Gln Leu Gly Arg Ile Asp 2100 2105 2110		
Val Val Gln Pro Val Leu Phe Ala Met Glu Val Ala Leu Ser Ala Leu 2115 2120 2125		
Trp Arg Ser Trp Gly Val Glu Pro Glu Ala Val Val Gly His Ser Met 2130 2135 2140		
Gly Glu Val Ala Ala Ala His Val Ala Gly Ala Leu Ser Leu Glu Asp 2145 2150 2155 2160		
Ala Val Ala Ile Ile Cys Arg Arg Ser Arg Leu Leu Arg Arg Ile Ser 2165 2170 2175		
Gly Gln Gly Glu Met Ala Leu Val Glu Leu Ser Leu Glu Glu Ala Glu 2180 2185 2190		
Ala Ala Leu Arg Gly His Glu Gly Arg Leu Ser Val Ala Val Ser Asn 2195 2200 2205		
Ser Pro Arg Ser Thr Val Leu Ala Gly Glu Pro Ala Ala Leu Ser Glu 2210 2215 2220		
Val Leu Ala Ala Leu Thr Ala Lys Gly Val Phe Trp Arg Gln Val Lys 2225 2230 2235 2240		
Val Asp Val Ala Ser His Ser Pro Gln Val Asp Pro Leu Arg Glu Glu 2245 2250 2255		
Leu Ile Ala Ala Leu Gly Ala Ile Arg Pro Arg Ala Ala Ala Val Pro 2260 2265 2270		
Met Arg Ser Thr Val Thr Gly Gly Val Ile Ala Gly Pro Glu Leu Gly 2275 2280 2285		
Ala Ser Tyr Trp Ala Asp Asn Leu Arg Gln Pro Val Arg Phe Ala Ala 2290 2295 2300		
Ala Ala Gln Ala Leu Leu Glu Gly Gly Pro Ala Leu Phe Ile Glu Met 2305 2310 2315 2320		
Ser Pro His Pro Ile Leu Val Pro Pro Leu Asp Glu Ile Gln Thr Ala 2325 2330 2335		
Ala Glu Gln Gly Gly Ala Ala Val Gly Ser Leu Arg Arg Gly Gln Asp 2340 2345 2350		
Glu Arg Ala Thr Leu Leu Glu Ala Leu Gly Thr Leu Trp Ala Ser Gly 2355 2360 2365		
Tyr Pro Val Ser Trp Ala Arg Leu Phe Pro Ala Gly Gly Arg Arg Val		

4

2725	2730	2735
Ser Val Val Tyr Arg Lys Val Ile Lys Arg Trp Met Glu His Leu Val 2740	2745	2750
Ala Ile Gly Ile Leu Val Gly Asp Gly Glu His Phe Val Ser Ser Gln 2755	2760	2765
Pro Leu Pro Glu Pro Asp Leu Ala Ala Val Leu Glu Glu Ala Gly Arg 2770	2775	2780
Val Phe Ala Asp Leu Pro Val Leu Phe Glu Trp Cys Lys Phe Ala Gly 2785	2790	2795
Glu Arg Leu Ala Asp Val Leu Thr Gly Lys Thr Leu Ala Leu Glu Ile 2805	2810	2815
Leu Phe Pro Gly Gly Ser Phe Asp Met Ala Glu Arg Ile Tyr Arg Asp 2820	2825	2830
Ser Pro Ile Ala Arg Tyr Ser Asn Gly Ile Val Arg Gly Val Val Glu 2835	2840	2845
Ser Ala Ala Arg Val Val Ala Pro Ser Gly Met Phe Ser Ile Leu Glu 2850	2855	2860
Ile Gly Ala Gly Thr Gly Ala Thr Thr Ala Ala Val Leu Pro Val Leu 2865	2870	2875
Leu Pro Asp Arg Thr Glu Tyr His Phe Thr Asp Val Ser Pro Leu Phe 2885	2890	2895
Leu Ala Arg Ala Glu Gln Arg Phe Arg Asp Tyr Pro Phe Leu Lys Tyr 2900	2905	2910
Gly Ile Leu Asp Val Asp Gln Glu Pro Ala Gly Gln Gly Tyr Ala His 2915	2920	2925
Gln Arg Phe Asp Val Ile Val Ala Ala Asn Val Ile His Ala Thr Arg 2930	2935	2940
Asp Ile Arg Ala Thr Ala Lys Arg Leu Leu Ser Leu Leu Ala Pro Gly 2945	2950	2955
Gly Leu Leu Val Leu Val Glu Gly Thr Gly His Pro Ile Trp Phe Asp 2965	2970	2975
Ile Thr Thr Gly Leu Ile Glu Gly Trp Gln Lys Tyr Glu Asp Asp Leu 2980	2985	2990
Arg Ile Asp His Pro Leu Leu Pro Ala Arg Thr Trp Cys Asp Val Leu 2995	3000	3005
Arg Arg Val Gly Phe Ala Asp Ala Val Ser Leu Pro Gly Asp Gly Ser 3010	3015	3020
Pro Ala Gly Ile Leu Gly Gln His Val Ile Leu Ser Arg Ala Pro Gly 3025	3030	3035
Ile Ala Gly Ala Ala Cys Asp Ser Ser Gly Glu Ser Ala Thr Glu Ser 3045	3050	3055
Pro Ala Ala Arg Ala Val Arg Gln Glu Trp Ala Asp Gly Ser Ala Asp 3060	3065	3070
Val Val His Arg Met Ala Leu Glu Arg Met Tyr Phe His Arg Arg Pro		

3075	3080	3085
Gly Arg Gln Val Trp Val His Gly Arg Leu Arg Thr Gly Gly Gly Ala 3090 3095 3100		
Phe Thr Lys Ala Leu Ala Gly Asp Leu Leu Leu Phe Glu Asp Thr Gly 3105 3110 3115 3120		
Gln Val Val Ala Glu Val Gln Gly Leu Arg Leu Pro Gln Leu Glu Ala 3125 3130 3135		
Ser Ala Phe Ala Pro Arg Asp Pro Arg Glu Glu Trp Leu Tyr Ala Leu 3140 3145 3150		
Glu Trp Gln Arg Lys Asp Pro Ile Pro Glu Ala Pro Ala Ala Ala Ser 3155 3160 3165		
Ser Ser Ser Ala Gly Ala Trp Leu Val Leu Met Asp Gln Gly Gly Thr 3170 3175 3180		
Gly Ala Ala Leu Val Ser Leu Leu Glu Gly Arg Gly Glu Ala Cys Val 3185 3190 3195 3200		
Arg Val Ile Ala Gly Thr Ala Tyr Ala Cys Leu Ala Pro Gly Leu Tyr 3205 3210 3215		
Gln Val Asp Pro Ala Gln Pro Asp Gly Phe His Thr Leu Leu Arg Asp 3220 3225 3230		
Ala Phe Gly Glu Asp Arg Ile Cys Arg Ala Val Val His Met Trp Ser 3235 3240 3245		
Leu Asp Ala Thr Ala Ala Gly Glu Arg Ala Thr Ala Glu Ser Leu Gln 3250 3255 3260		
Ala Asp Gln Leu Leu Gly Ser Leu Ser Ala Leu Ser Leu Val Gln Ala 3265 3270 3275 3280		
Leu Val Arg Arg Arg Trp Arg Asn Met Pro Arg Leu Trp Leu Leu Thr 3285 3290 3295		
Arg Ala Val His Ala Val Gly Ala Glu Asp Ala Ala Ala Ser Val Ala 3300 3305 3310		
Gln Ala Pro Val Trp Gly Leu Gly Arg Thr Leu Ala Leu Glu His Pro 3315 3320 3325		
Glu Leu Arg Cys Thr Leu Val Asp Val Asn Pro Ala Pro Ser Pro Glu 3330 3335 3340		
Asp Ala Ala Ala Leu Ala Val Glu Leu Gly Ala Ser Asp Arg Glu Asp 3345 3350 3355 3360		
Gln Val Ala Leu Arg Ser Asp Gly Arg Tyr Val Ala Arg Leu Val Arg 3365 3370 3375		
Ser Ser Phe Ser Gly Lys Pro Ala Thr Asp Cys Gly Ile Arg Ala Asp 3380 3385 3390		
Gly Ser Tyr Val Ile Thr Asp Gly Met Gly Arg Val Gly Leu Ser Val 3395 3400 3405		
Ala Gln Trp Met Val Met Gln Gly Ala Arg His Val Val Leu Val Asp 3410 3415 3420		
Arg Gly Gly Ala Ser Glu Ala Ser Arg Asp Ala Leu Arg Ser Met Ala		

3425	3430	3435	3440
Glu Ala Gly	Ala Glu Val Gln Ile Val Glu Ala Asp Val Ala Arg Arg		
	3445	3450	3455
Asp Asp Val	Ala Arg Leu Leu Ser Lys Ile Glu Pro Ser Met Pro Pro		
	3460	3465	3470
Leu Arg Gly	Ile Val Tyr Val Asp Gly Thr Phe Gln Gly Asp Ser Ser		
	3475	3480	3485
Met Leu Glu	Leu Asp Ala Arg Arg Phe Lys Glu Trp Met Tyr Pro Lys		
	3490	3495	3500
Val Leu Gly	Ala Trp Asn Leu His Ala Leu Thr Arg Asp Arg Ser Leu		
	3505	3510	3515
Asp Phe Phe	Val Leu Tyr Ser Ser Gly Thr Ser Leu Leu Gly Leu Pro		
	3525	3530	3535
Gly Gln Gly	Ser Arg Ala Ala Gly Asp Ala Phe Leu Asp Ala Ile Ala		
	3540	3545	3550
His His Arg	Cys Lys Val Gly Leu Thr Ala Met Ser Ile Asn Trp Gly		
	3555	3560	3565
Leu Leu Ser	Glu Ala Ser Ser Pro Ala Thr Pro Asn Asp Gly Gly Ala		
	3570	3575	3580
Arg Leu Glu	Tyr Arg Gly Met Glu Gly Leu Thr Leu Glu Gln Gly Ala		
	3585	3590	3595
Ala Ala Leu	Gly Arg Leu Leu Ala Arg Pro Arg Ala Gln Val Gly Val		
	3605	3610	3615
Met Arg Leu	Asn Leu Arg Gln Trp Leu Glu Phe Tyr Pro Asn Ala Ala		
	3620	3625	3630
Arg Leu Ala	Leu Trp Ala Glu Leu Leu Lys Glu Arg Asp Arg Ala Asp		
	3635	3640	3645
Arg Gly Ala	Ser Asn Ala Ser Asn Leu Arg Glu Ala Leu Gln Ser Ala		
	3650	3655	3660
Arg Pro Glu	Asp Arg Gln Leu Ile Leu Glu Lys His Leu Ser Glu Leu		
	3665	3670	3675
Leu Gly Arg	Gly Leu Arg Leu Pro Pro Glu Arg Ile Glu Arg His Val		
	3685	3690	3695
Pro Phe Ser	Asn Leu Gly Met Asp Ser Leu Ile Gly Leu Glu Leu Arg		
	3700	3705	3710
Asn Arg Ile	Glu Ala Ala Leu Gly Ile Thr Val Pro Ala Thr Leu Leu		
	3715	3720	3725
Trp Thr Tyr	Pro Asn Val Ala Ala Leu Ser Gly Ser Leu Leu Asp Ile		
	3730	3735	3740
Leu Phe Pro	Asn Ala Gly Ala Thr His Ala Pro Ala Thr Glu Arg Glu		
	3745	3750	3755
Lys Ser Phe	Glu Asn Asp Ala Ala Asp Leu Glu Ala Leu Arg Gly Met		
	3765	3770	3775
Thr Asp Glu	Gln Lys Asp Ala Leu Leu Ala Glu Lys Leu Ala Gln Leu		



3780

3785

3790

Ala Gln Ile Val Gly Glu  
3795

&lt;210&gt; 7

&lt;211&gt; 2439

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 7

Met Ala Thr Thr Asn Ala Gly Lys Leu Glu His Ala Leu Leu Leu Met  
1 5 10 15

Asp Lys Leu Ala Lys Lys Asn Ala Ser Leu Glu Gln Glu Arg Thr Glu  
20 25 30

Pro Ile Ala Ile Val Gly Ile Gly Cys Arg Phe Pro Gly Gly Ala Asp  
35 40 45

Thr Pro Glu Ala Phe Trp Glu Leu Leu Asp Ser Gly Arg Asp Ala Val  
50 55 60

Gln Pro Leu Asp Arg Arg Trp Ala Leu Val Gly Val His Pro Ser Glu  
65 70 75 80

Glu Val Pro Arg Trp Ala Gly Leu Leu Thr Glu Ala Val Asp Gly Phe  
85 90 95

Asp Ala Ala Phe Phe Gly Thr Ser Pro Arg Glu Ala Arg Ser Leu Asp  
100 105 110

Pro Gln Gln Arg Leu Leu Leu Glu Val Thr Trp Glu Gly Leu Glu Asp  
115 120 125

Ala Gly Ile Ala Pro Gln Ser Leu Asp Gly Ser Arg Thr Gly Val Phe  
130 135 140

Leu Gly Ala Cys Ser Ser Asp Tyr Ser His Thr Val Ala Gln Gln Arg  
145 150 155 160

Arg Glu Glu Gln Asp Ala Tyr Asp Ile Thr Gly Asn Thr Leu Ser Val  
165 170 175

Ala Ala Gly Arg Leu Ser Tyr Thr Leu Gly Leu Gln Gly Pro Cys Leu  
180 185 190

Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Ile His Leu Ala  
195 200 205

Cys Arg Ser Leu Arg Ala Arg Glu Ser Asp Leu Ala Leu Ala Gly Gly  
210 215 220

Val Asn Met Leu Leu Ser Ser Lys Thr Met Ile Met Leu Gly Arg Ile  
225 230 235 240

Gln Ala Leu Ser Pro Asp Gly His Cys Arg Thr Phe Asp Ala Ser Ala  
245 250 255

Asn Gly Phe Val Arg Gly Glu Gly Cys Gly Met Val Val Leu Lys Arg  
260 265 270

Leu Ser Asp Ala Gln Arg His Gly Asp Arg Ile Trp Ala Leu Ile Arg  
275 280 285

Gly Ser Ala Met Asn Gln Asp Gly Arg Ser Thr Gly Leu Met Ala Pro  
 290 295 300  
 Asn Val Leu Ala Gln Glu Ala Leu Leu Arg Glu Ala Leu Gln Ser Ala  
 305 310 315 320  
 Arg Val Asp Ala Gly Ala Ile Gly Tyr Val Glu Thr His Gly Thr Gly  
 325 330 335  
 Thr Ser Leu Gly Asp Pro Ile Glu Val Glu Ala Leu Arg Ala Val Leu  
 340 345 350  
 Gly Pro Ala Arg Ala Asp Gly Ser Arg Cys Val Leu Gly Ala Val Lys  
 355 360 365  
 Thr Asn Leu Gly His Leu Glu Gly Ala Ala Gly Val Ala Gly Leu Ile  
 370 375 380  
 Lys Ala Ala Leu Ala Leu His His Glu Leu Ile Pro Arg Asn Leu His  
 385 390 395 400  
 Phe His Thr Leu Asn Pro Arg Ile Arg Ile Glu Gly Thr Ala Leu Ala  
 405 410 415  
 Leu Ala Thr Glu Pro Val Pro Trp Pro Arg Ala Gly Arg Pro Arg Phe  
 420 425 430  
 Ala Gly Val Ser Ala Phe Gly Leu Ser Gly Thr Asn Val His Val Val  
 435 440 445  
 Leu Glu Glu Ala Pro Ala Thr Val Leu Ala Pro Ala Thr Pro Gly Arg  
 450 455 460  
 Ser Ala Glu Leu Leu Val Leu Ser Ala Lys Ser Ala Ala Ala Leu Asp  
 465 470 475 480  
 Ala Gln Ala Ala Arg Leu Ser Ala His Ile Ala Ala Tyr Pro Glu Gln  
 485 490 495  
 Gly Leu Gly Asp Val Ala Phe Ser Leu Val Ser Thr Arg Ser Pro Met  
 500 505 510  
 Glu His Arg Leu Ala Val Ala Ala Thr Ser Arg Glu Ala Leu Arg Ser  
 515 520 525  
 Ala Leu Glu Val Ala Ala Gln Gly Gln Thr Pro Ala Gly Ala Ala Arg  
 530 535 540  
 Gly Arg Ala Ala Ser Ser Pro Gly Lys Leu Ala Phe Leu Phe Ala Gly  
 545 550 555 560  
 Gln Gly Ala Gln Val Pro Gly Met Gly Arg Gly Leu Trp Glu Ala Trp  
 565 570 575  
 Pro Ala Phe Arg Glu Thr Phe Asp Arg Cys Val Thr Leu Phe Asp Arg  
 580 585 590  
 Glu Leu His Gln Pro Leu Cys Glu Val Met Trp Ala Glu Pro Gly Ser  
 595 600 605  
 Ser Arg Ser Ser Leu Leu Asp Gln Thr Ala Phe Thr Gln Pro Ala Leu  
 610 615 620  
 Phe Ala Leu Glu Tyr Ala Leu Ala Ala Leu Phe Arg Ser Trp Gly Val  
 625 630 635 640

Glu Pro Glu Leu Val Ala Gly His Ser Leu Gly Glu Leu Val Ala Ala  
 645 650 655  
 Cys Val Ala Gly Val Phe Ser Leu Glu Asp Ala Val Arg Leu Val Val  
 660 665 670  
 Ala Arg Gly Arg Leu Met Gln Ala Leu Pro Ala Gly Gly Ala Met Val  
 675 680 685  
 Ser Ile Ala Ala Pro Glu Ala Asp Val Ala Ala Ala Val Ala Pro His  
 690 695 700  
 Ala Ala Leu Val Ser Ile Ala Ala Val Asn Gly Pro Glu Gln Val Val  
 705 710 715 720  
 Ile Ala Gly Ala Glu Lys Phe Val Gln Gln Ile Ala Ala Ala Phe Ala  
 725 730 735  
 Ala Arg Gly Ala Arg Thr Lys Pro Leu His Val Ser His Ala Phe His  
 740 745 750  
 Ser Pro Leu Met Asp Pro Met Leu Glu Ala Phe Arg Arg Val Thr Glu  
 755 760 765  
 Ser Val Thr Tyr Arg Arg Pro Ser Ile Ala Leu Val Ser Asn Leu Ser  
 770 775 780  
 Gly Lys Pro Cys Thr Asp Glu Val Ser Ala Pro Gly Tyr Trp Val Arg  
 785 790 795 800  
 His Ala Arg Glu Ala Val Arg Phe Ala Asp Gly Val Lys Ala Leu His  
 805 810 815  
 Ala Ala Gly Ala Gly Leu Phe Val Glu Val Gly Pro Lys Pro Thr Leu  
 820 825 830  
 Leu Gly Leu Val Pro Ala Cys Leu Pro Asp Ala Arg Pro Val Leu Leu  
 835 840 845  
 Pro Ala Ser Arg Ala Gly Arg Asp Glu Ala Ala Ser Ala Leu Glu Ala  
 850 855 860  
 Leu Gly Gly Phe Trp Val Val Gly Gly Ser Val Thr Trp Ser Gly Val  
 865 870 875 880  
 Phe Pro Ser Gly Gly Arg Arg Val Pro Leu Pro Thr Tyr Pro Trp Gln  
 885 890 895  
 Arg Glu Arg Tyr Trp Ile Glu Ala Pro Val Asp Arg Glu Ala Asp Gly  
 900 905 910  
 Thr Gly Arg Ala Arg Ala Gly Gly His Pro Leu Leu Gly Glu Val Phe  
 915 920 925  
 Ser Val Ser Thr His Ala Gly Leu Arg Leu Trp Glu Thr Thr Leu Asp  
 930 935 940  
 Arg Lys Arg Leu Pro Trp Leu Gly Glu His Arg Ala Gln Gly Glu Val  
 945 950 955 960  
 Val Phe Pro Gly Ala Gly Tyr Leu Glu Met Ala Leu Ser Ser Gly Ala  
 965 970 975  
 Glu Ile Leu Gly Asp Gly Pro Ile Gln Val Thr Asp Val Val Leu Ile  
 980 985 990

- 67 -

Glu Thr Leu Thr Phe Ala Gly Asp Thr Ala Val Pro Val Gln Val Val  
 995 1000 1005  
 Thr Thr Glu Glu Arg Pro Gly Arg Leu Arg Phe Gln Val Ala Ser Arg  
 1010 1015 1020  
 Glu Pro Gly Glu Arg Arg Ala Pro Phe Arg Ile His Ala Arg Gly Val  
 1025 1030 1035 1040  
 Leu Arg Arg Ile Gly Arg Val Glu Thr Pro Ala Arg Ser Asn Leu Ala  
 1045 1050 1055  
 Ala Leu Arg Ala Arg Leu His Ala Ala Val Pro Ala Ala Ala Ile Tyr  
 1060 1065 1070  
 Gly Ala Leu Ala Glu Met Gly Leu Gln Tyr Gly Pro Ala Leu Arg Gly  
 1075 1080 1085  
 Leu Ala Glu Leu Trp Arg Gly Glu Gly Glu Ala Leu Gly Arg Val Arg  
 1090 1095 1100  
 Leu Pro Glu Ala Ala Gly Ser Ala Thr Ala Tyr Gln Leu His Pro Val  
 1105 1110 1115 1120  
 Leu Leu Asp Ala Cys Val Gln Met Ile Val Gly Ala Phe Ala Asp Arg  
 1125 1130 1135  
 Asp Glu Ala Thr Pro Trp Ala Pro Val Glu Val Gly Ser Val Arg Leu  
 1140 1145 1150  
 Phe Gln Arg Ser Pro Gly Glu Leu Trp Cys His Ala Arg Val Val Ser  
 1155 1160 1165  
 Asp Gly Gln Gln Ala Ser Ser Arg Trp Ser Ala Asp Phe Glu Leu Met  
 1170 1175 1180  
 Asp Gly Thr Gly Ala Val Val Ala Glu Ile Ser Arg Leu Val Val Glu  
 1185 1190 1195 1200  
 Arg Leu Ala Ser Gly Val Arg Arg Arg Asp Ala Asp Asp Trp Phe Leu  
 1205 1210 1215  
 Glu Leu Asp Trp Glu Pro Ala Ala Leu Gly Gly Pro Lys Ile Thr Ala  
 1220 1225 1230  
 Gly Arg Trp Leu Leu Leu Gly Glu Gly Gly Gly Leu Gly Arg Ser Leu  
 1235 1240 1245  
 Cys Ser Ala Leu Lys Ala Ala Gly His Val Val Val His Ala Ala Gly  
 1250 1255 1260  
 Asp Asp Thr Ser Thr Ala Gly Met Arg Ala Leu Leu Ala Asn Ala Phe  
 1265 1270 1275 1280  
 Asp Gly Gln Ala Pro Thr Ala Val Val His Leu Ser Ser Leu Asp Gly  
 1285 1290 1295  
 Gly Gly Gln Leu Gly Pro Gly Leu Gly Ala Gln Gly Ala Leu Asp Ala  
 1300 1305 1310  
 Pro Arg Ser Pro Asp Val Asp Ala Asp Ala Leu Glu Ser Ala Leu Met  
 1315 1320 1325  
 Arg Gly Cys Asp Ser Val Leu Ser Leu Val Gln Ala Leu Val Gly Met  
 1330 1335 1340

- 68 -

Asp Leu Arg Asn Ala Pro Arg Leu Trp Leu Leu Thr Arg Gly Ala Gln  
 1345 1350 1355 1360  
 Ala Ala Ala Ala Gly Asp Val Ser Val Val Gln Ala Pro Leu Leu Gly  
 1365 1370 1375  
 Leu Gly Arg Thr Ile Ala Leu Glu His Ala Glu Leu Arg Cys Ile Ser  
 1380 1385 1390  
 Val Asp Leu Asp Pro Ala Glu Pro Glu Gly Glu Ala Asp Ala Leu Leu  
 1395 1400 1405  
 Ala Glu Leu Leu Ala Asp Asp Ala Glu Glu Glu Val Ala Leu Arg Gly  
 1410 1415 1420  
 Gly Asp Arg Leu Val Ala Arg Leu Val His Arg Leu Pro Asp Ala Gln  
 1425 1430 1435 1440  
 Arg Arg Glu Lys Val Glu Pro Ala Gly Asp Arg Pro Phe Arg Leu Glu  
 1445 1450 1455  
 Ile Asp Glu Pro Gly Ala Leu Asp Gln Leu Val Leu Arg Ala Thr Gly  
 1460 1465 1470  
 Arg Arg Ala Pro Gly Pro Gly Glu Val Glu Ile Ser Val Glu Ala Ala  
 1475 1480 1485  
 Gly Leu Asp Ser Ile Asp Ile Gln Leu Ala Leu Gly Val Ala Pro Asn  
 1490 1495 1500  
 Asp Leu Pro Gly Glu Glu Ile Glu Pro Leu Val Leu Gly Ser Glu Cys  
 1505 1510 1515 1520  
 Ala Gly Arg Ile Val Ala Val Gly Glu Gly Val Asn Gly Leu Val Val  
 1525 1530 1535  
 Gly Gln Pro Val Ile Ala Leu Ala Ala Gly Val Phe Ala Thr His Val  
 1540 1545 1550  
 Thr Thr Ser Ala Thr Leu Val Leu Pro Arg Pro Leu Gly Leu Ser Ala  
 1555 1560 1565  
 Thr Glu Ala Ala Ala Met Pro Leu Ala Tyr Leu Thr Ala Trp Tyr Ala  
 1570 1575 1580  
 Leu Asp Lys Val Ala His Leu Gln Ala Gly Glu Arg Val Leu Ile His  
 1585 1590 1595 1600  
 Ala Glu Ala Gly Gly Val Gly Leu Cys Ala Val Arg Trp Ala Gln Arg  
 1605 1610 1615  
 Val Gly Ala Glu Val Tyr Ala Thr Ala Asp Thr Pro Glu Asn Arg Ala  
 1620 1625 1630  
 Tyr Leu Glu Ser Leu Gly Val Arg Tyr Val Ser Asp Ser Arg Ser Gly  
 1635 1640 1645  
 Arg Phe Val Thr Asp Val His Ala Trp Thr Asp Gly Glu Gly Val Asp  
 1650 1655 1660  
 Val Val Leu Asp Ser Leu Ser Gly Glu Arg Ile Asp Lys Ser Leu Met  
 1665 1670 1675 1680  
 Val Leu Arg Ala Cys Gly Arg Leu Val Lys Leu Gly Arg Arg Asp Asp  
 1685 1690 1695

Cys Ala Asp Thr Gln Pro Gly Leu Pro Pro Leu Leu Arg Asn Phe Ser  
 1700 1705 1710  
 Phe Ser Gln Val Asp Leu Arg Gly Met Met Leu Asp Gln Pro Ala Arg  
 1715 1720 1725  
 Ile Arg Ala Leu Leu Asp Glu Leu Phe Gly Leu Val Ala Ala Gly Ala  
 1730 1735 1740  
 Ile Ser Pro Leu Gly Ser Gly Leu Arg Val Gly Gly Ser Leu Thr Pro  
 1745 1750 1755 1760  
 Pro Pro Val Glu Thr Phe Pro Ile Ser Arg Ala Ala Glu Ala Phe Arg  
 1765 1770 1775  
 Arg Met Ala Gln Gly Gln His Leu Gly Lys Leu Val Leu Thr Leu Asp  
 1780 1785 1790  
 Asp Pro Glu Val Arg Ile Arg Ala Pro Ala Glu Ser Ser Val Ala Val  
 1795 1800 1805  
 Arg Ala Asp Gly Thr Tyr Leu Val Thr Gly Gly Leu Gly Gly Leu Gly  
 1810 1815 1820  
 Leu Arg Val Ala Gly Trp Leu Ala Glu Arg Gly Ala Gly Gln Leu Val  
 1825 1830 1835 1840  
 Leu Val Gly Arg Ser Gly Ala Ala Ser Ala Glu Gln Arg Ala Ala Val  
 1845 1850 1855  
 Ala Ala Leu Glu Ala His Gly Ala Arg Val Thr Val Ala Lys Ala Asp  
 1860 1865 1870  
 Val Ala Asp Arg Ser Gln Ile Glu Arg Val Leu Arg Glu Val Thr Ala  
 1875 1880 1885  
 Ser Gly Met Pro Leu Arg Gly Val Val His Ala Ala Gly Leu Val Asp  
 1890 1895 1900  
 Asp Gly Leu Leu Met Gln Gln Thr Pro Ala Arg Phe Arg Thr Val Met  
 1905 1910 1915 1920  
 Gly Pro Lys Val Gln Gly Ala Leu His Leu His Thr Leu Thr Arg Glu  
 1925 1930 1935  
 Ala Pro Leu Ser Phe Phe Val Leu Tyr Ala Ser Ala Ala Gly Leu Phe  
 1940 1945 1950  
 Gly Ser Pro Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Phe Leu Asp  
 1955 1960 1965  
 Ala Leu Ser His His Arg Arg Ala Gln Gly Leu Pro Ala Leu Ser Ile  
 1970 1975 1980  
 Asp Trp Gly Met Phe Thr Glu Val Gly Met Ala Val Ala Gln Glu Asn  
 1985 1990 1995 2000  
 Arg Gly Ala Arg Gln Ile Ser Arg Gly Met Arg Gly Ile Thr Pro Asp  
 2005 2010 2015  
 Glu Gly Leu Ser Ala Leu Ala Arg Leu Leu Glu Gly Asp Arg Val Gln  
 2020 2025 2030  
 Thr Gly Val Ile Pro Ile Thr Pro Arg Gln Trp Val Glu Phe Tyr Pro  
 2035 2040 2045

Ala Thr Ala Ala Ser Arg Arg Leu Ser Arg Leu Val Thr Thr Gln Arg  
 2050 2055 2060  
 Ala Val Ala Asp Arg Thr Ala Gly Asp Arg Asp Leu Leu Glu Gln Leu  
 2065 2070 2075 2080  
 Ala Ser Ala Glu Pro Ser Ala Arg Ala Gly Leu Leu Gln Asp Val Val  
 2085 2090 2095  
 Arg Val Gln Val Ser His Val Leu Arg Leu Pro Glu Asp Lys Ile Glu  
 2100 2105 2110  
 Val Asp Ala Pro Leu Ser Ser Met Gly Met Asp Ser Leu Met Ser Leu  
 2115 2120 2125  
 Glu Leu Arg Asn Arg Ile Glu Ala Ala Leu Gly Val Ala Ala Pro Ala  
 2130 2135 2140  
 Ala Leu Gly Trp Thr Tyr Pro Thr Val Ala Ala Ile Thr Arg Trp Leu  
 2145 2150 2155 2160  
 Leu Asp Asp Ala Leu Val Val Arg Leu Gly Gly Gly Ser Asp Thr Asp  
 2165 2170 2175  
 Glu Ser Thr Ala Ser Ala Gly Ser Phe Val His Val Leu Arg Phe Arg  
 2180 2185 2190  
 Pro Val Val Lys Pro Arg Ala Arg Leu Phe Cys Phe His Gly Ser Gly  
 2195 2200 2205  
 Gly Ser Pro Glu Gly Phe Arg Ser Trp Ser Glu Lys Ser Glu Trp Ser  
 2210 2215 2220  
 Asp Leu Glu Ile Val Ala Met Trp His Asp Arg Ser Leu Ala Ser Glu  
 2225 2230 2235 2240  
 Asp Ala Pro Gly Lys Lys Tyr Val Gln Glu Ala Ala Ser Leu Ile Gln  
 2245 2250 2255  
 His Tyr Ala Asp Ala Pro Phe Ala Leu Val Gly Phe Ser Leu Gly Val  
 2260 2265 2270  
 Arg Phe Val Met Gly Thr Ala Val Glu Leu Ala Ser Arg Ser Gly Ala  
 2275 2280 2285  
 Pro Ala Pro Leu Ala Val Phe Thr Leu Gly Gly Ser Leu Ile Ser Ser  
 2290 2295 2300  
 Ser Glu Ile Thr Pro Glu Met Glu Thr Asp Ile Ile Ala Lys Leu Phe  
 2305 2310 2315 2320  
 Phe Arg Asn Ala Ala Gly Phe Val Arg Ser Thr Gln Gln Val Gln Ala  
 2325 2330 2335  
 Asp Ala Arg Ala Asp Lys Val Ile Thr Asp Thr Met Val Ala Pro Ala  
 2340 2345 2350  
 Pro Gly Asp Ser Lys Glu Pro Pro Val Lys Ile Ala Val Pro Ile Val  
 2355 2360 2365  
 Ala Ile Ala Gly Ser Asp Asp Val Ile Val Pro Pro Ser Asp Val Gln  
 2370 2375 2380  
 Asp Leu Gln Ser Arg Thr Thr Glu Arg Phe Tyr Met His Leu Leu Pro  
 2385 2390 2395 2400

Asp Thr Thr Ile Tyr Leu Ile Ala Phe Ala Val Leu Asn Leu Leu Arg  
260 265 270



Ser Pro Glu Ala Leu Glu Leu Val Lys Ala Glu Pro Gly Leu Met Arg  
275 280 285

Asn Ala Leu Asp Glu Val Leu Arg Phe Asp Asn Ile Leu Arg Ile Gly  
290 295 300

Thr Val Arg Phe Ala Arg Gln Asp Leu Glu Tyr Cys Gly Ala Ser Ile  
305 310 315 320

Lys Lys Gly Glu Met Val Phe Leu Leu Ile Pro Ser Ala Leu Arg Asp  
325 330 335

Gly Thr Val Phe Ser Arg Pro Asp Val Phe Asp Val Arg Arg Asp Thr  
340 345 350

Gly Ala Ser Leu Ala Tyr Gly Arg Gly Pro His Val Cys Pro Gly Val  
355 360 365

Ser Leu Ala Arg Leu Glu Ala Glu Ile Ala Val Gly Thr Ile Phe Arg  
370 375 380

Arg Phe Pro Glu Met Lys Leu Lys Glu Thr Pro Val Phe Gly Tyr His  
385 390 395 400

Pro Ala Phe Arg Asn Ile Glu Ser Leu Asn Val Ile Leu Lys Pro Ser  
405 410 415

Lys Ala Gly

<210> 9

<211> 607

<212> PRT

<213> Sorangium cellulosum

<400> 9

Ala Ser Leu Asp Ala Leu Phe Ala Arg Ala Thr Ser Ala Arg Val Leu  
1 5 10 15

Asp Asp Gly His Gly Arg Ala Thr Glu Arg His Val Leu Ala Glu Ala  
20 25 30

Arg Gly Ile Glu Asp Leu Arg Ala Leu Arg Glu His Leu Arg Ile Gln  
35 40 45

Glu Gly Gly Pro Ser Phe His Cys Met Cys Leu Gly Asp Leu Thr Val  
50 55 60

Glu Leu Leu Ala His Asp Gln Pro Leu Ala Ser Ile Ser Phe His His  
65 70 75 80

Ala Arg Ser Leu Arg His Pro Asp Trp Thr Ser Asp Ala Met Leu Val  
85 90 95

Asp Gly Pro Ala Leu Val Arg Trp Leu Ala Ala Arg Gly Ala Pro Gly  
100 105 110

Pro Leu Arg Glu Tyr Glu Glu Glu Arg Glu Arg Ala Arg Thr Ala Gln  
115 120 125

Glu Ala Arg Arg Leu Trp Leu Ala Ala Ala Pro Pro Cys Phe Ala Pro  
130 135 140

Asp Leu Pro Arg Phe Glu Asp Asp Ala Asn Gly Leu Pro Leu Gly Pro

145                      150                      155                      160  
 Met Ser Pro Glu Val Ala Glu Ala Glu Arg Arg Leu Arg Ala Ser Tyr  
                                  165                                   170                                   175  
 Ala Thr Pro Glu Leu Ala Cys Ala Ala Leu Leu Ala Trp Leu Gly Thr  
                                  180                                   185                                   190  
 Gly Ala Gly Pro Trp Ser Gly Tyr Pro Ala Tyr Glu Met Leu Pro Glu  
                                  195                                   200                                   205  
 Asn Leu Leu Leu Gly Phe Gly Leu Pro Thr Ala Ile Ala Ala Ala Ser  
                                  210                                   215                                   220  
 Ala Pro Gly Thr Ser Glu Ala Ala Leu Arg Gly Ala Ala Arg Leu Phe  
                                  225                                   230                                   235                                   240  
 Ala Ser Trp Glu Val Val Ser Ser Lys Lys Ser Gln Leu Gly Asn Ile  
    245                                   250                                   255  
 Pro Glu Ala Leu Trp Glu Arg Leu Arg Thr Ile Val Arg Ala Met Gly  
    260                                   265                                   270  
 Asn Ala Asp Asn Leu Ser Arg Phe Glu Arg Ala Glu Ala Ile Ala Ala  
    275                                   280                                   285  
 Glu Val Arg Arg Leu Arg Ala Gln Pro Ala Pro Phe Ala Ala Gly Ala  
    290                                   295                                   300  
 Gly Leu Ala Val Ala Gly Val Ser Ser Ser Gly Arg Leu Ser Gly Leu  
    305                                   310                                   315                                   320  
 Val Thr Asp Gly Asp Ala Leu Tyr Ser Gly Asp Gly Asn Asp Ile Val  
    325                                   330                                   335  
 Met Phe Gln Pro Gly Arg Ile Ser Pro Val Val Leu Leu Ala Gly Thr  
    340                                   345                                   350  
 Asp Pro Phe Phe Glu Leu Ala Pro Pro Leu Ser Gln Met Leu Phe Val  
    355                                   360                                   365  
 Ala His Ala Asn Ala Gly Thr Ile Ser Lys Val Leu Thr Glu Gly Ser  
    370                                   375                                   380  
 Pro Leu Ile Val Met Ala Arg Asn Gln Ala Arg Pro Met Ser Leu Val  
    385                                   390                                   395                                   400  
 His Ala Arg Gly Phe Met Ala Trp Val Asn Gln Ala Met Val Pro Asp  
    405                                   410                                   415  
 Pro Glu Arg Gly Ala Pro Phe Val Val Gln Arg Ser Thr Ile Met Glu  
    420                                   425                                   430  
 Phe Glu His Pro Thr Pro Arg Cys Leu His Glu Pro Ala Gly Ser Ala  
    435                                   440                                   445  
 Phe Ser Leu Ala Cys Asp Glu Glu His Leu Tyr Trp Cys Glu Leu Ser  
    450                                   455                                   460  
 Ala Gly Arg Leu Glu Leu Trp Arg His Pro His His Arg Pro Gly Ala  
    465                                   470                                   475                                   480  
 Pro Ser Arg Phe Ala Tyr Leu Gly Glu His Pro Ile Ala Ala Thr Trp  
    485                                   490                                   495  
 Tyr Pro Ser Leu Thr Leu Asn Ala Thr His Val Leu Trp Ala Asp Pro

500										505					510						
Asp	Arg	Arg	Ala	Ile	Leu	Gly	Val	Asp	Lys	Arg	Thr	Gly	Val	Glu	Pro						
		515					520					525									
Ile	Val	Leu	Ala	Glu	Thr	Arg	His	Pro	Pro	Ala	His	Val	Val	Ser	Glu						
		530				535					540										
Asp	Arg	Asp	Ile	Phe	Ala	Leu	Thr	Gly	Gln	Pro	Asp	Ser	Arg	Asp	Trp						
545				550						555					560						
His	Val	Glu	His	Ile	Arg	Ser	Gly	Ala	Ser	Thr	Val	Val	Ala	Asp	Tyr						
			565						570					575							
Gln	Arg	Gln	Leu	Trp	Asp	Arg	Pro	Asp	Met	Val	Leu	Asn	Arg	Arg	Gly						
			580					585					590								
Leu	Phe	Phe	Thr	Thr	Asn	Asp	Arg	Ile	Leu	Thr	Leu	Ala	Arg	Ser							
		595					600					605									

&lt;210&gt; 10

&lt;211&gt; 423

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 10

Met	Gly	Ala	Leu	Ile	Ser	Val	Ala	Ala	Pro	Gly	Cys	Ala	Leu	Gly	Gly
1				5					10					15	

Ala	Glu	Glu	Glu	Gly	Gln	Pro	Gly	Gln	Asp	Ala	Gly	Ala	Gly	Ala	Leu
		20						25					30		

Ala	Pro	Ala	Arg	Glu	Val	Met	Ala	Ala	Glu	Val	Ala	Ala	Gly	Gln	Met
		35					40					45			

Pro	Gly	Ala	Val	Trp	Leu	Val	Ala	Arg	Gly	Asp	Asp	Val	His	Val	Asp
	50					55					60				

Ala	Val	Gly	Val	Thr	Glu	Leu	Gly	Gly	Ser	Ala	Pro	Met	Arg	Arg	Asp
	65				70					75					80

Thr	Ile	Phe	Arg	Ile	Ala	Ser	Met	Thr	Lys	Ala	Val	Thr	Ala	Thr	Ala
				85					90					95	

Val	Met	Met	Leu	Val	Glu	Glu	Gly	Lys	Leu	Asp	Leu	Asp	Ser	Pro	Val
			100					105					110		

Asp	Arg	Trp	Leu	Pro	Glu	Leu	Ala	Asn	Arg	Lys	Val	Leu	Ala	Arg	Ile
		115					120					125			

Asp	Gly	Pro	Ile	Asp	Glu	Thr	Val	Pro	Ala	Glu	Arg	Pro	Ile	Thr	Val
	130					135					140				

Arg	Asp	Leu	Met	Thr	Phe	Thr	Met	Gly	Phe	Gly	Ile	Ser	Phe	Asp	Ala
	145				150					155					160

Ser	Ser	Pro	Ile	Gln	Arg	Ala	Ile	Asp	Glu	Leu	Gly	Leu	Val	Asn	Ala
				165					170					175	

Gln	Pro	Val	Pro	Met	Thr	Pro	His	Gly	Pro	Asp	Glu	Trp	Ile	Arg	Arg
			180					185					190		

Leu	Gly	Thr	Leu	Pro	Leu	Met	His	Gln	Pro	Gly	Ala	Gln	Trp	Met	Tyr
		195					200					205			

- 75 -

Asn Thr Gly Ser Leu Val Gln Gly Val Leu Val Gly Arg Ala Ala Asp  
210 215 220

Gln Gly Phe Asp Ala Phe Val Arg Glu Arg Ile Leu Ala Pro Leu Gly  
225 230 235 240

Met Arg Asp Thr Asp Phe His Val Pro Ala Asp Lys Leu Ala Arg Phe  
245 250 255

Ala Gly Cys Gly Tyr Phe Thr Asp Glu Gln Thr Gly Glu Lys Thr Arg  
260 265 270

Met Asp Arg Asp Gly Ala Glu Ser Ala Tyr Ala Ser Pro Pro Ala Phe  
275 280 285

Pro Ser Gly Ala Ala Gly Leu Val Ser Thr Val Asp Asp Tyr Leu Leu  
290 295 300

Phe Ala Arg Met Leu Met Asn Gly Gly Val His Glu Gly Arg Arg Leu  
305 310 315 320

Leu Ser Ala Ala Ser Val Arg Glu Met Thr Ala Asp His Leu Thr Pro  
325 330 335

Ala Gln Lys Ala Ala Ser Ser Phe Phe Pro Gly Phe Phe Glu Thr His  
340 345 350

Gly Trp Gly Tyr Gly Met Ala Val Val Thr Ala Pro Asp Ala Val Ser  
355 360 365

Glu Val Pro Gly Arg Tyr Gly Trp Asp Gly Gly Phe Gly Thr Ser Trp  
370 375 380

Ile Asn Asp Pro Gly Arg Glu Leu Ile Gly Ile Val Met Thr Gln Ser  
385 390 395 400

Ala Gly Phe Leu Phe Ser Gly Ala Leu Glu Arg Phe Trp Arg Ser Val  
405 410 415

Tyr Val Ala Thr Glu Ser Ala  
420

<210> 11

<211> 713

<212> PRT

<213> Sorangium cellulosum

<400> 11

Met His Gly Leu Thr Glu Arg Gln Val Leu Leu Ser Leu Val Thr Leu  
1 5 10 15

Ala Leu Ile Leu Val Thr Ala Arg Ala Ser Gly Glu Leu Ala Arg Arg  
20 25 30

Leu Arg Gln Pro Glu Val Leu Gly Glu Leu Phe Gly Gly Val Val Leu  
35 40 45

Gly Pro Ser Val Val Gly Ala Leu Ala Pro Gly Phe His Arg Ala Leu  
50 55 60

Phe Gln Glu Pro Ala Val Gly Val Val Leu Ser Gly Ile Ser Trp Ile  
65 70 75 80

Gly Ala Leu Leu Leu Leu Met Ala Gly Ile Glu Val Asp Val Gly  
85 90 95

Ile Leu Arg Lys Glu Ala Arg Pro Gly Ala Leu Ser Ala Leu Gly Ala  
 100 105 110  
 Ile Ala Pro Pro Leu Ala Ala Gly Ala Ala Phe Ser Ala Leu Val Leu  
 115 120 125  
 Asp Arg Pro Leu Pro Ser Gly Leu Phe Leu Gly Ile Val Leu Ser Val  
 130 135 140  
 Thr Ala Val Ser Val Ile Ala Lys Val Leu Ile Glu Arg Glu Ser Met  
 145 150 155 160  
 Arg Arg Ser Tyr Ala Gln Val Thr Leu Ala Ala Gly Val Val Ser Glu  
 165 170 175  
 Val Ala Ala Trp Val Leu Val Ala Met Thr Ser Ser Ser Tyr Gly Ala  
 180 185 190  
 Ser Pro Ala Leu Ala Val Ala Arg Ser Ala Leu Leu Ala Ser Gly Phe  
 195 200 205  
 Leu Leu Phe Met Val Leu Val Gly Arg Arg Leu Thr His Leu Ala Met  
 210 215 220  
 Arg Trp Val Ala Asp Ala Thr Arg Val Ser Lys Gly Gln Val Ser Leu  
 225 230 235 240  
 Val Leu Val Leu Thr Phe Leu Ala Ala Ala Leu Thr Gln Arg Leu Gly  
 245 250 255  
 Leu His Pro Leu Leu Gly Ala Phe Ala Leu Gly Val Leu Leu Asn Ser  
 260 265 270  
 Ala Pro Arg Thr Asn Arg Pro Leu Leu Asp Gly Val Gln Thr Leu Val  
 275 280 285  
 Ala Gly Leu Phe Ala Pro Val Phe Phe Val Leu Ala Gly Met Arg Val  
 290 295 300  
 Asp Val Ser Gln Leu Arg Thr Pro Ala Ala Trp Gly Thr Val Ala Leu  
 305 310 315 320  
 Leu Leu Ala Thr Ala Thr Ala Ala Lys Val Val Pro Ala Ala Leu Gly  
 325 330 335  
 Ala Arg Leu Gly Gly Leu Arg Gly Ser Glu Ala Ala Leu Val Ala Val  
 340 345 350  
 Gly Leu Asn Met Lys Gly Gly Thr Asp Leu Ile Val Ala Ile Val Gly  
 355 360 365  
 Val Glu Leu Gly Leu Leu Ser Asn Glu Ala Tyr Thr Met Tyr Ala Val  
 370 375 380  
 Val Ala Leu Val Thr Val Thr Ala Ser Pro Ala Leu Leu Ile Trp Leu  
 385 390 395 400  
 Glu Lys Arg Ala Pro Pro Thr Gln Glu Glu Ser Ala Arg Leu Glu Arg  
 405 410 415  
 Glu Glu Ala Ala Arg Arg Ala Tyr Ile Pro Gly Val Glu Arg Ile Leu  
 420 425 430  
 Val Pro Ile Val Ala His Ala Leu Pro Gly Phe Ala Thr Asp Ile Val  
 435 440 445

- 77 -

Glu Ser Ile Val Ala Ser Lys Arg Lys Leu Gly Glu Thr Val Asp Ile  
 450 455 460  
 Thr Glu Leu Ser Val Glu Gln Gln Ala Pro Gly Pro Ser Arg Ala Ala  
 465 470 475 480  
 Gly Glu Ala Ser Arg Gly Leu Ala Arg Leu Gly Ala Arg Leu Arg Val  
 485 490 495  
 Gly Ile Trp Arg Gln Arg Arg Glu Leu Arg Gly Ser Ile Gln Ala Ile  
 500 505 510  
 Leu Arg Ala Ser Arg Asp His Asp Leu Leu Val Ile Gly Ala Arg Ser  
 515 520 525  
 Pro Ala Arg Ala Arg Gly Met Ser Phe Gly Arg Leu Gln Asp Ala Ile  
 530 535 540  
 Val Gln Arg Ala Glu Ser Asn Val Leu Val Val Val Gly Asp Pro Pro  
 545 550 555 560  
 Ala Ala Glu Arg Ala Ser Ala Arg Arg Ile Leu Val Pro Ile Ile Gly  
 565 570 575  
 Leu Glu Tyr Ser Phe Ala Ala Ala Asp Leu Ala Ala His Val Ala Leu  
 580 585 590  
 Ala Trp Asp Ala Glu Leu Val Leu Leu Ser Ser Ala Gln Thr Asp Pro  
 595 600 605  
 Gly Ala Val Val Trp Arg Asp Arg Glu Pro Ser Arg Val Arg Ala Val  
 610 615 620  
 Ala Arg Ser Val Val Asp Glu Ala Val Phe Arg Gly Arg Arg Leu Gly  
 625 630 635 640  
 Val Arg Val Ser Ser Arg Val His Val Gly Ala His Pro Ser Asp Glu  
 645 650 655  
 Ile Thr Arg Glu Leu Ala Arg Ala Pro Tyr Asp Leu Leu Val Leu Gly  
 660 665 670  
 Cys Tyr Asp His Gly Pro Leu Gly Arg Leu Tyr Leu Gly Ser Thr Val  
 675 680 685  
 Glu Ser Val Val Val Arg Ser Arg Val Pro Val Ala Leu Leu Val Ala  
 690 695 700  
 His Gly Gly Thr Arg Glu Gln Val Arg  
 705 710

<210> 12  
 <211> 126  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 12  
 Met Asp Lys Pro Ile Gly Arg Thr Arg Cys Ala Ile Ala Glu Gly Tyr  
 1 5 10 15  
 Ile Pro Gly Gly Ser Asn Gly Pro Glu Pro Gln Met Thr Ser His Glu  
 20 25 30  
 Thr Ala Cys Leu Leu Asn Ala Ser Asp Arg Asp Ala Gln Val Ala Ile

- 78 -

35                      40                      45  
 Thr Val Tyr Phe Ser Asp Arg Asp Pro Ala Gly Pro Tyr Arg Val Thr  
     50                      55                      60  
 Val Pro Ala Arg Arg Thr Arg His Val Arg Phe Asn Asp Leu Thr Glu  
     65                      70                      75                      80  
 Pro Glu Pro Ile Pro Arg Asp Thr Asp Tyr Ala Ser Val Ile Glu Ser  
                     85                      90                      95  
 Asp Ala Pro Ile Val Val Gln His Thr Arg Leu Asp Ser Arg Gln Ala  
                     100                      105                      110  
 Glu Asn Ala Leu Leu Ser Thr Ile Ala Tyr Thr Asp Arg Glu  
                     115                      120                      125

<210> 13  
 <211> 149  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 13  
 Met Lys His Val Asp Thr Gly Arg Arg Phe Gly Arg Arg Ile Gly His  
     1                      5                      10                      15  
 Thr Leu Gly Leu Leu Ala Ser Met Ala Leu Ala Gly Cys Gly Gly Pro  
                     20                      25                      30  
 Ser Glu Lys Thr Val Gln Gly Thr Arg Leu Ala Pro Gly Ala Asp Ala  
                     35                      40                      45  
 Arg Val Thr Ala Asp Val Asp Pro Asp Ala Ala Thr Thr Arg Leu Ala  
                     50                      55                      60  
 Val Asp Val Val His Leu Ser Pro Pro Glu Arg Leu Glu Ala Gly Ser  
     65                      70                      75                      80  
 Glu Arg Phe Val Val Trp Gln Arg Pro Ser Pro Glu Ser Pro Trp Arg  
                     85                      90                      95  
 Arg Val Gly Val Leu Asp Tyr Asn Ala Asp Ser Arg Arg Gly Lys Leu  
                     100                      105                      110  
 Ala Glu Thr Thr Val Pro Tyr Ala Asn Phe Glu Leu Leu Ile Thr Ala  
                     115                      120                      125  
 Glu Lys Gln Ser Ser Pro Gln Ser Pro Ser Ser Ala Ala Val Ile Gly  
                     130                      135                      140  
 Pro Thr Ser Val Gly  
 145

<210> 14  
 <211> 184  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 14  
 Val Thr Ser Glu Glu Val Pro Gly Ala Ala Leu Gly Ala Gln Ser Ser  
     1                      5                      10                      15  
 Leu Val Arg Ala Gln His Ala Ala Arg His Val Arg Pro Cys Thr Arg  
                     20                      25                      30

Ala Glu Glu Pro Pro Ala Leu Met His Gly Leu Thr Glu Arg Gln Val  
                   35                                  40                                  45

Leu Leu Ser Leu Val Ala Leu Ala Leu Val Leu Leu Thr Ala Arg Ala  
           50                                  55                                  60

Phe Gly Glu Leu Ala Arg Arg Leu Arg Gln Pro Glu Val Leu Gly Glu  
       65                                  70                                  75                                  80

Leu Phe Gly Gly Val Val Leu Gly Pro Ser Val Val Gly Ala Leu Ala  
                                   85                                  90                                  95

Pro Gly Phe His Arg Val Leu Phe Gln Asp Pro Ala Val Gly Val Val  
                                   100                                  105                                  110

Leu Ser Gly Ile Ser Trp Ile Gly Ala Leu Val Leu Leu Leu Met Ala  
                   115                                  120                                  125

Gly Ile Glu Val Asp Val Ser Ile Leu Arg Lys Glu Ala Arg Pro Gly  
       130                                  135                                  140

Ala Leu Ser Ala Leu Gly Ala Ile Ala Pro Pro Leu Arg Thr Pro Gly  
       145                                  150                                  155                                  160

Pro Leu Val Gln Arg Met Gln Gly Ala Phe Thr Trp Asp Leu Asp Val  
                                   165                                  170                                  175

Ser Pro Arg Arg Ser Ala Gln Ala  
                                   180

<210> 15  
 <211> 145  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 15  
 Val Asn Ala Pro Cys Met Arg Cys Thr Ser Gly Pro Gly Val Arg Ser  
   1                                  5                                  10                                  15

Gly Gly Ala Ile Ala Pro Ser Ala Glu Ser Ala Pro Gly Arg Ala Ser  
                   20                                  25                                  30

Leu Arg Arg Met Leu Thr Ser Thr Ser Ile Pro Ala Met Ser Ser Arg  
                   35                                  40                                  45

Thr Ser Ala Pro Ile Gln Glu Met Pro Glu Ser Thr Thr Pro Thr Ala  
           50                                  55                                  60

Gly Ser Trp Lys Arg Thr Arg Trp Asn Pro Gly Ala Ser Ala Pro Thr  
       65                                  70                                  75                                  80

Thr Asp Gly Pro Ser Thr Thr Pro Pro Lys Ser Ser Pro Ser Thr Ser  
                                   85                                  90                                  95

Gly Trp Arg Ser Arg Arg Ala Ser Ser Pro Lys Ala Arg Ala Val Arg  
                   100                                  105                                  110

Arg Thr Ser Ala Arg Ala Thr Ser Glu Ser Arg Thr Cys Arg Ser Val  
                   115                                  120                                  125

Arg Pro Cys Ile Arg Ala Gly Gly Ser Ser Ala Arg Val Gln Gly Arg  
       130                                  135                                  140

Thr



145

&lt;210&gt; 16

&lt;211&gt; 185

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 16

Val Leu Ala Pro Pro Ala Asp Ile Arg Pro Pro Ala Ala Ala Gln Leu  
 1 5 10 15

Glu Pro Asp Ser Pro Asp Asp Glu Ala Asp Glu Ala Asp Glu Ala Leu  
 20 25 30

Arg Pro Phe Arg Asp Ala Ile Ala Ala Tyr Ser Glu Ala Val Arg Trp  
 35 40 45

Ala Glu Ala Ala Gln Arg Pro Arg Leu Glu Ser Leu Val Arg Leu Ala  
 50 55 60

Ile Val Arg Leu Gly Lys Ala Leu Asp Lys Val Pro Phe Ala His Thr  
 65 70 75 80

Thr Ala Gly Val Ser Gln Ile Ala Gly Arg Leu Gln Asn Asp Ala Val  
 85 90 95

Trp Phe Asp Val Ala Ala Arg Tyr Ala Ser Phe Arg Ala Ala Thr Glu  
 100 105 110

His Ala Leu Arg Asp Ala Ala Ser Ala Met Glu Ala Leu Ala Ala Gly  
 115 120 125

Pro Tyr Arg Gly Ser Ser Arg Val Ser Ala Ala Val Gly Glu Phe Arg  
 130 135 140

Gly Glu Ala Ala Arg Leu His Pro Ala Asp Arg Val Pro Ala Ser Asp  
 145 150 155 160

Gln Gln Ile Leu Thr Ala Leu Arg Ala Ala Glu Arg Ala Leu Ile Ala  
 165 170 175

Leu Tyr Thr Ala Phe Ala Arg Glu Glu  
 180 185

&lt;210&gt; 17

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Sorangium cellulosum

&lt;400&gt; 17

Met Ala Asp Ala Ala Ser Arg Ser Ala Cys Ser Val Ala Ala Arg Lys  
 1 5 10 15

Leu Ala Tyr Arg Ala Ala Thr Ser Asn Gln Thr Ala Ser Phe Trp Ser  
 20 25 30

Leu Pro Ala Ile Trp Glu Thr Pro Ala Val Val Cys Ala Lys Gly Thr  
 35 40 45

Leu Ser Ser Ala Leu Pro Ser Arg Thr Ile Ala Ser Arg Thr Arg Leu  
 50 55 60

Ser Ser Arg Gly Arg Cys Ala Ala Ser Ala His Arg Thr Ala Ser Glu  
 65 70 75 80

[illegible]

```
<210> 18
<211> 288
<212> PRT
<213> Sorangium cellulosum
```

<400>	18																
Val	Thr	Val	Ser	Ser	Met	Pro	Arg	Ser	Trp	Ser	Ser	Arg	Val	Arg	Thr		
1				5					10					15			
Val	Val	Thr	Ala	Leu	Gly	Cys	Ala	Arg	Arg	Leu	Ser	Gly	Ser	Ile	Ser		
			20					25					30				
Arg	Leu	Arg	Arg	His	Pro	Glu	Ala	Gly	Arg	Ala	Pro	Arg	Ser	Arg	Leu		
		35					40					45					
Arg	Ala	Trp	Arg	Arg	Leu	Pro	Gln	His	Ile	Ser	Ser	Pro	Trp	Arg	His		
	50					55					60						
Leu	Pro	Pro	Gly	Ala	Arg	Val	Gly	Thr	Ser	Cys	Pro	Ala	Asp	Arg	Arg		
	65				70					75					80		
Ile	Leu	Pro	Ser	His	Arg	Thr	Ala	Asp	Leu	Gly	Thr	Ser	Gly	Gly	Thr		
				85					90					95			
Leu	Val	Ala	Arg	Met	Ser	Gly	His	Val	Ala	Arg	Asn	Pro	His	Ala	Ala		
			100					105					110				
Val	Leu	Val	Gly	Asp	Gly	Ser	Ala	Arg	Gly	Arg	Arg	Arg	Leu	Ser	Asn		
		115					120					125					
Arg	Arg	Ala	Glu	Arg	Arg	Val	Ser	Asp	Val	Thr	Cys	Arg	Glu	Gly	Gly		
		130				135					140						
Glu	Ala	Met	Gln	Lys	Ile	Ala	Gly	Lys	Leu	Val	Val	Gly	Leu	Ile	Ser		
	145				150					155					160		
Val	Ser	Gly	Met	Ser	Leu	Leu	Ala	Ala	Cys	Gly	Gly	Glu	Lys	Arg	Ser		
				165					170					175			
Gly	Gly	Glu	Ala	Gln	Thr	Pro	Gly	Gly	Ala	Gln	Gly	Glu	Ala	Pro	Val		
			180					185					190				
Pro	Val	Gly	Ser	Ala	Val	Asp	Ser	Ile	Val	Ala	Ala	Arg	Cys	Asp	Arg		
		195					200					205					
Glu	Ala	Arg	Cys	Asn	Asn	Ile	Gly	Gln	Asp	Arg	Glu	Tyr	Ser	Ser	Lys		
	210					215					220						
Asp	Ala	Cys	Ser	Asn	Lys	Ile	Arg	Ser	Glu	Trp	Arg	Asp	Glu	Leu	Thr		

- 82 -

225	230	235	240
Phe Gly Glu Cys	Pro Gly Gly Ile Asp	Ala Lys Gln Leu Asn	Glu Cys
	245	250	255
Leu Glu Gly Ile	Arg Asn Glu Gly Cys	Gly Asn Pro Phe Asp	Thr Leu
	260	265	270
Gly Arg Val Val	Ala Cys Arg Ser Ser	Asp Leu Cys Arg	Asp Ala Arg
	275	280	285

<210> 19  
 <211> 288  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 19  
 Val Thr Val Ser Ser Met Pro Arg Ser Trp Ser Ser Arg Val Arg Thr  
 1 5 10 15  
 Val Val Thr Ala Leu Gly Cys Ala Arg Arg Leu Ser Gly Ser Ile Ser  
 20 25 30  
 Arg Leu Arg Arg His Pro Glu Ala Gly Arg Ala Pro Arg Ser Arg Leu  
 35 40 45  
 Arg Ala Trp Arg Arg Leu Pro Gln His Ile Ser Ser Pro Trp Arg His  
 50 55 60  
 Leu Pro Pro Gly Ala Arg Val Gly Thr Ser Cys Pro Ala Asp Arg Arg  
 65 70 75 80  
 Ile Leu Pro Ser His Arg Thr Ala Asp Leu Gly Thr Ser Gly Gly Thr  
 85 90 95  
 Leu Val Ala Arg Met Ser Gly His Val Ala Arg Asn Pro His Ala Ala  
 100 105 110  
 Val Leu Val Gly Asp Gly Ser Ala Arg Gly Arg Arg Arg Leu Ser Asn  
 115 120 125  
 Arg Arg Ala Glu Arg Arg Val Ser Asp Val Thr Cys Arg Glu Gly Gly  
 130 135 140  
 Glu Ala Met Gln Lys Ile Ala Gly Lys Leu Val Val Gly Leu Ile Ser  
 145 150 155 160  
 Val Ser Gly Met Ser Leu Leu Ala Ala Cys Gly Gly Glu Lys Arg Ser  
 165 170 175  
 Gly Gly Glu Ala Gln Thr Pro Gly Gly Ala Gln Gly Glu Ala Pro Val  
 180 185 190  
 Pro Val Gly Ser Ala Val Asp Ser Ile Val Ala Ala Arg Cys Asp Arg  
 195 200 205  
 Glu Ala Arg Cys Asn Asn Ile Gly Gln Asp Arg Glu Tyr Ser Ser Lys  
 210 215 220  
 Asp Ala Cys Ser Asn Lys Ile Arg Ser Glu Trp Arg Asp Glu Leu Thr  
 225 230 235 240

Phe Gly Glu Cys Pro Gly Gly Ile Asp Ala Lys Gln Leu Asn Glu Cys  
245 250 255

Leu Glu Gly Ile Arg Asn Glu Gly Cys Gly Asn Pro Phe Asp Thr Leu  
260 265 270

Gly Arg Val Val Ala Cys Arg Ser Ser Asp Leu Cys Arg Asp Ala Arg  
275 280 285

```
<210> 20
<211> 155
<212> PRT
<213> Sorangium cellulosum
```

```

<400> 20
Met Asp Pro Arg Ala Arg Arg Glu Lys Arg Pro Ser Leu Leu Asp Ser
  1          5          10          15
Arg Gly Arg Gln Pro Lys Arg Ser Gln Gln Gly Gly His Met Glu Lys
          20          25          30
Pro Ile Gly Arg Thr Arg Trp Ala Ile Ala Glu Gly Tyr Ile Pro Gly
      35          40          45
Arg Ser Asn Gly Pro Glu Pro Gln Met Thr Ser His Glu Thr Ala Cys
  50          55          60
Leu Leu Asn Ala Ser Asp Arg Asp Ala Gln Val Ala Ile Thr Val Tyr
  65          70          75          80
Phe Ser Asp Arg Asp Pro Ala Gly Pro Tyr Arg Val Thr Val Pro Ala
          85          90          95
Arg Arg Thr Arg His Val Arg Phe Asn Asp Leu Thr Glu Pro Glu Pro
          100          105          110
Ile Pro Arg Asp Thr Asp Tyr Ala Ser Val Ile Glu Ser Asp Val Pro
      115          120          125
Ile Val Val Gln His Thr Arg Leu Asp Ser Arg Gln Ala Glu Asn Ala
      130          135          140
Leu Ile Ser Thr Ile Ala Tyr Thr Asp Arg Glu
145          150          155

```

```
<210> 21
<211> 156
<212> PRT
<213> Sorangium cellulosum
```

<400> 21  
Val Arg Arg Ser Arg Trp Gln Met Lys His Val Asp Thr Gly Arg Arg  
1 5 10 15  
Val Gly Arg Arg Ile Gly Leu Thr Leu Gly Leu Leu Ala Ser Met Ala  
20 25 30  
Leu Ala Gly Cys Gly Gly Pro Ser Glu Lys Ile Val Gln Gly Thr Arg  
35 40 45  
Leu Ala Pro Gly Ala Asp Ala His Val Ala Ala Asp Val Asp Pro Asp

50

55

60

Ala Ala Thr Thr Arg Leu Ala Val Asp Val Val His Leu Ser Pro Pro  
 65 70 75 80  
 Glu Arg Ile Glu Ala Gly Ser Glu Arg Phe Val Val Trp Gln Arg Pro  
 85 90 95  
 Ser Ser Glu Ser Pro Trp Gln Arg Val Gly Val Leu Asp Tyr Asn Ala  
 100 105 110  
 Ala Ser Arg Arg Gly Lys Leu Ala Glu Thr Thr Val Pro His Ala Asn  
 115 120 125  
 Phe Glu Leu Leu Ile Thr Val Glu Lys Gln Ser Ser Pro Gln Ser Pro  
 130 135 140  
 Ser Ser Ala Ala Val Ile Gly Pro Thr Ser Val Gly  
 145 150 155

<210> 22  
 <211> 305  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 22  
 Met Glu Lys Glu Ser Arg Ile Ala Ile Tyr Gly Ala Ile Ala Ala Asn  
 1 5 10 15  
 Val Ala Ile Ala Ala Val Lys Phe Ile Ala Ala Ala Val Thr Gly Ser  
 20 25 30  
 Ser Ala Met Leu Ser Glu Gly Val His Ser Leu Val Asp Thr Ala Asp  
 35 40 45  
 Gly Leu Leu Leu Leu Leu Gly Lys His Arg Ser Ala Arg Pro Pro Asp  
 50 55 60  
 Ala Glu His Pro Phe Gly His Gly Lys Glu Leu Tyr Phe Trp Thr Leu  
 65 70 75 80  
 Ile Val Ala Ile Met Ile Phe Ala Ala Gly Gly Gly Val Ser Ile Tyr  
 85 90 95  
 Glu Gly Ile Leu His Leu Leu His Pro Arg Gln Ile Glu Asp Pro Thr  
 100 105 110  
 Trp Asn Tyr Val Val Leu Gly Ala Ala Ala Val Phe Glu Gly Thr Ser  
 115 120 125  
 Leu Ile Ile Ser Ile His Glu Phe Lys Lys Lys Asp Gly Gln Gly Tyr  
 130 135 140  
 Leu Ala Ala Met Arg Ser Ser Lys Asp Pro Thr Thr Phe Thr Ile Val  
 145 150 155 160  
 Leu Glu Asp Ser Ala Ala Leu Ala Gly Leu Thr Ile Ala Phe Leu Gly  
 165 170 175  
 Val Trp Leu Gly His Arg Leu Gly Asn Pro Tyr Leu Asp Gly Ala Ala  
 180 185 190  
 Ser Ile Gly Ile Gly Leu Val Leu Ala Ala Val Ala Val Phe Leu Ala  
 195 200 205

- 85 -

Ser Gln Ser Arg Gly Leu Leu Val Gly Glu Ser Ala Asp Arg Glu Leu  
 210 215 220  
 Leu Ala Ala Ile Arg Ala Leu Ala Ser Ala Asp Pro Gly Val Ser Ala  
 225 230 235 240  
 Val Gly Arg Pro Leu Thr Met His Phe Gly Pro His Glu Val Leu Val  
 245 250 255  
 Val Leu Arg Ile Glu Phe Asp Ala Ala Leu Thr Ala Ser Gly Val Ala  
 260 265 270  
 Glu Ala Ile Glu Arg Ile Glu Thr Arg Ile Arg Ser Glu Arg Pro Asp  
 275 280 285  
 Val Lys His Ile Tyr Val Glu Ala Arg Ser Leu His Gln Arg Ala Arg  
 290 295 300

Ala  
 305

<210> 23  
 <211> 135  
 <212> PRT  
 <213> Sorangium cellulosum

<400> 23  
 Val Gln Thr Ser Ser Phe Asp Ala Arg Tyr Ala Gly Cys Lys Ser Ser  
 1 5 10 15  
 Arg Arg Ile Ala Arg Ser Gly Ser Ala Gly Ala Arg Ala Gly Arg Ala  
 20 25 30  
 His Glu Gly Ala Ala Ser Ala Gly Phe Glu Gly Gly Asp Val Met Arg  
 35 40 45  
 Lys Ala Arg Ala His Gly Ala Met Leu Gly Gly Arg Asp Asp Gly Trp  
 50 55 60  
 Arg Arg Gly Leu Pro Gly Ala Gly Ala Leu Arg Ala Ala Leu Gln Arg  
 65 70 75 80  
 Gly Arg Ser Arg Asp Leu Ala Arg Arg Arg Leu Ile Ala Ser Val Ser  
 85 90 95  
 Leu Ala Gly Gly Ala Ser Met Ala Val Val Ser Leu Phe Gln Leu Gly  
 100 105 110  
 Ile Ile Glu Arg Leu Pro Asp Pro Pro Leu Pro Gly Phe Asp Ser Ala  
 115 120 125  
 Lys Val Thr Ser Ser Asp Ile  
 130 135

<210> 24  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: universal  
 reverse primer

<400> 24

ggaaacagct atgaccatg

19

<210> 25

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: universal  
forward primer

<400> 25

gtaaaacgac ggccagt

17

<210> 26

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer  
NH24 end "B"

<400> 26

gtgactggcg cctggaatct gcatgagc

28

<210> 27

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer NH2  
end "A"

<400> 27

agcgggagct tgctagacat tctgtttc

28

<210> 28

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer NH2  
end "B"

<400> 28

gacgcgcctc gggcagcgcc ccaa

24

<210> 29

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer  
pEPO15-NH6 end "B"

<400> 29

caccgaagcg tcgatctggt ccatc

25

<210> 30

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer  
pEP015H2.7 end "A"

<400> 30

cggtcagatc gacgacgggc ttcc

25